Development of pea albumin based GNA fusion proteins as novel biopesticides for the control of aphid (Hemiptera: aphidae) crop pests.

DE-THIER, JAKE,SIMON

Development of pea albumin based GNA fusion proteins as novel biopesticides for the control of aphid (Hemiptera: aphidae) crop pests.

A thesis submitted by Jake Simon De Thier, BSc in accordance with the requirements of Durham University for the degree of Doctor of Philosophy

Department of Biosciences

Durham University

September 2019

Abstract

Development of pea albumin based GNA fusion proteins as novel biopesticides for the control of aphid (Hemiptera: aphidae) crop pests.

Jake Simon De Thier

Novel insecticidal compounds with new modes of action are urgently required as current reliance and overuse of chemical pesticides that target a limited number of sites within insect pests has driven widespread development of resistance in crop pests. Furthermore, the most widely utilised biopesticide *Bacillus thuringiensis* (Bt) targets lepidopteran, coleopteran and certain dipteran pests but is ineffective against hemipteran sap-sucking pests. Hemipteran sap-sucking pests cause considerable damage to crops through direct damage and transmission of many economically important plant viruses. Further, they have been documented to rapidly develop resistance to novel insecticides. Fusion of invertebrate specific neurotoxins to snowdrop lectin (*Galanthus nivalis* agglutinin; GNA) has previously been shown to enhance oral activity of peptide toxins by virtue of its ability to bind to the insect gut epithelium and subsequently to transport attached peptides across the gut to the circulatory system. In addition to delivering attached toxins, GNA via binding to insect gut polypeptides, was hypothesised to provide an opportunity to potentiate the efficacy of insecticidal proteins that act at the gut surface. Pea albumin (PA1b), from pea (*Pisum sativum*) seeds is a proteinaceous inhibitor of V-ATPase proton pumps that are localised to the insect gut epithelium. The PA1 gene cassette is expressed *in planta* as a preproprotein and cleaved into two mature proteins, PA1b and PA1a. PA1b has previously been shown to cause mortality when supplied in diet to several insect species from different orders including hemipteran pest species. PA1a has no documented insecticidal activity. This project focuses on the development and production of novel recombinant fusion proteins based on linking PA1b to GNA for the control of three different aphid species (pea aphid; *Acrythosiphon pisum*, peach potato aphid; *Myzus persicae* and grain aphid; *Sitobion avenae*).

Production of a highly pure recombinant pea albumin – GNA fusion protein (PA1b/GNA) in *Pichia pastoris* proved challenging; multiple expression constructs with varying linker regions, orientations and histidine purification tags were typically expressed at relatively low levels and recovery from nickel affinity purification columns was poor. However, the use of a fusion protein expression construct that included a strep-tactin affinity tag allowed the recovery of relatively pure protein samples from fermented *P. pastoris* cultures. An LC_{50} of 7 nmol mL^{-1} PA1b/GNA against *A. pisum* was subsequently derived from artificial diet bioassays and this suggested that fusion of PA1b to GNA did indeed enhance toxicity as compared to activity reported for PA1b

alone in the literature. However, due to low recovery rates for PA1b/GNA, an alternative approach to generate sufficient quantities of highly pure fusion proteins based upon constructs encoding both PA1b and PA1a linked to GNA (PAF/GNA) was pursued. Expression of PAF/GNA in *P. pastoris* resulted in higher expression levels, and more importantly, allowed the recovery of greater quantities of highly pure protein samples as compared to PA1b/GNA.

The oral toxicity of PAF/GNA and recombinant PAF was assessed through artificial diet bioassays against three species of aphid pests: *A. pisum* (LC_{50} = 5 nmol mL^{-1}), *S. avenae* (15 nmol mL^{-1}) and *M. persicae* (34 nmol mL^{-1}). PAF/GNA was significantly more orally toxic towards *A. pisum* and *M. persicae* as compared to recombinant PAF or GNA alone. Greater retention of PAF/GNA in the *A. pisum* gut over time, as compared to controls (PAF, GNA or ovalbumin) was demonstrated using confocal microscopy of *A. pisum* aphids pulse fed on fluorescently labelled proteins. Significant differences in toxicity were observed between *A. pisum* and *M. persicae*, it was hypothesised that this may be due to the stability of PAF/GNA when exposed to insect proteases. PAF/GNA, GNA and PAF were all similarly stable when exposed to salivary secretions from either species for up to 48 hours. Feeding assays were performed to assess the *in vivo* stability, no intact protein was detected in either species, but the quantity of breakdown products decreased in *A. pisum* over the 48 hours, possibly suggesting that the quantity of diet consumed was decreased compared to *M. persicae*. Antifeedant effects were indeed indicated in free choice experiments, *A. pisum* displayed a clear preference for diets without pea albumin or GNA, whereas *M. persicae* showed no significant preference except after 48 hours feeding on PAF/GNA where they preferentially fed on control diet.

Whilst PAF/GNA due to its potentially limited host range is unlikely to be utilised in a commercial setting, there may be a case made for the expression of a PAF/GNA fusion protein under the control of a phloem specific promoter as a novel method to control aphids that are one of the most economically damaging groups of agricultural pests.

Table of Contents

List of figures

List of tables

List of abbreviations

Abbreviation	Meaning
	Units of Measurement
L	Litre
mL	millilitre
µl	microlitre
nl	nanolitre
t	tons
kg	kilogram
g	gram
mg	milligram
µg	microgram
ng	nanogram
kDa	Kilodalton
M	Molar
mM	millimolar
µM	micromolar
nM	nanomolar
nmol	nanomoles
°C	Degrees Celsius
	Proteins, chemicals and biological molecules
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BMGY	Buffered Glycerol-complex Medium
BSA	Bovine Serum Albumin
BSM	Basal salt medium
Cry	Crystal toxin
Cyt	Cytolytic
DNA	Deoxyribose nucleic acid
FITC	Fluorescein isothiocyanate
GNA	*Galanthus nivalis* agglutinin
ICK	Inhibitor cysteine knot
LSLB	Low salt lysogeny broth

PAF	Pa1b+PA1a
PAF/GNA	PAF+GNA fusion protein
PBS	Phosphate buffered saline
PTM1 salts	*Pichia* trace metal 1 salts
RNA	Ribose nucleic acid
SDS	Sodium dodecyl sulphate
TFA	Trifluoroacetic acid
YPG	Yeast peptone glycerol medium

Amino Acids

Letter	Abbreviation	Amino Acid
A	Ala	Alanine
R	Arg	Arginine
N	Asn	Asparagine
D	Asp	Aspartic acid
C	Cys	Cysteine
Q	Gln	Glutamine
E	Glu	Glutamic acid
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
L	Leu	Leucine
K	Lys	Lysine
M	Met	Methionine
F	Phe	Phenylalanine
P	Pro	Proline
S	Ser	Serine
T	Thr	Threonine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
V	Val	Valine

Miscellaneous Abbreviations

Bt	*Bacillus thuringiensis*
BBSRC	Biotechnology and biological sciences research council
Bt	*Bacillus thuringiensis*
BYDD	Barley yellow dwarf disease
BYDD	Barley yellow dwarf disease
BYDV	Barley yellow dwarf virus
BYDV	Barley yellow dwarf virus
EU	European Union
EU	European Union
FAO	Food and Agricultural Organisation of the United Nations
FAO	Food and Agricultural Organisation of the United Nations
FT	follow-through
FT	follow-through
LC_{50}	Median lethal concentration
LC_{50}	Median lethal concentration
PCR	Polymerase chain reaction
PLRV	Potato leaf roll virus
SDS-PAGE	Sodium Dodecyl Sulphate – PolyAcrylamide Gel Electrophoresis
SUP	Supernatant
USA	United States of America
v	volume
w	weight

Declaration and statement of copyright

Declaration:

No material contained herein has been submitted for any other degree. Except where acknowledged, all material is the work of the author.

Statement of Copyright:

Date: ……..

Signed: ……..

Acknowledgements

First and foremost, I would like to thank my supervisor Dr Elaine Fitches (Department of Biosciences, Durham University) for her advice, support, guidance and unending patience throughout the course of my PhD studies. Without her guidance this work would not have been possible

The BBSRC is gratefully acknowledged for funding this work. I would also like to acknowledge Dr Prashant Pyati for his previous work on PA1b/GNA. I would also like to thank Prof John Gatehouse (Department of Biosciences, Durham University) for his insight throughout this project.

The members of Lab 1 past and present all deserve recognition for creating a friendly and fun working environment. I would particularly like to recognise Dr Prashant Pyati, Dr Min Cao for their help, advice and support during my PhD.

I am indebted to friends and family around the world who have made this journey possible through their support and encouragement, not just during the course of my PhD but throughout my entire education. It goes without saying that I have the deepest gratitude toward my parents, Annette and David De Thier, and step-parents, Andrew Beattie and Patricia De Thier, for their love and encouragement throughout this process. In particular I would like to thank, Dr Fieka Sukiran, Tracey Stephenson, and Jeannette Alcaraz for the laughs and company in the department over these last few years. Finally, and by no means least I'd like to thank Madeleine Harris, her love and support throughout most of this process has kept me going particularly during these last few months.

Jake De Thier

Chapter 1 |Introduction

1.1 General introduction

Currently there are more than 2 billion people on planet earth who do not have regular access to safe and nutritious food, 819 million people suffer from hunger and are considered undernourished (FAO *et al*., 2019). Asia and Africa, are particularly effected with 10 and 20 % of their populations undernourished, these percentages represent upwards of 700 million people (FAO *et al*., 2019). Couple these statistics with predictions from the United Nations (UN) which state that the global population will reach 9.7 billion people by 2050 and that most of this growth will occur in Africa and Asia and the challenges which need to be overcome to achieve global food security seem monumental (United Nations, 2019). The FAO have estimated that a 70 % increase in agricultural food production will need to occur in order to meet the increased demand, whilst crop yields are generally predicted to decrease in developing regions as a result of climate change (FAO, 2013, 2016). However, increases in productivity on this scale will be challenging due to several factors including soil degradation, limits in arable land, water availability, insect and disease damage and climate change (FAO, 2011). According to Tilman *et al*., (2011) in order to reach that level of production a further one billion hectares of land, an area larger than China, would need to be farmed and this would have significant negative impacts on biodiversity and many natural habitats. This is obviously not feasible and therefore improvement to agricultural technology is necessary.

The green revolution of the 1960s saw a period of rapid change in the agricultural landscape. The introduction of novel technologies, such as high yielding dwarf crop varieties and the expansion existing of technologies such as pesticides, synthetic fertilisers and modern irrigation systems to developing nations facilitated a 2.6 fold increase in global grain production (Kendall and Pimentel, 1994; Khush, 2001). Without many of these technologies losses can reach up to 80 % in sugar beet and cotton, and upwards of 50 % in barely (Oerke and Dehne, 2004). In order to meet the growing demand for food outlined above a second green revolution is required.

Insect herbivory is currently a major challenge to agriculture and causes a significant loss in crop yield. Fourteen percent of crops are lost annually to insect pests, which represents an economic loss of $244 billion per annum across eight major crops (Ferry and Gatehouse, 2010). Insects from multiple orders, including Hemiptera, Diptera, Coleoptera, Lepidoptera and Hymenoptera, act as crop pests cause significant damage through consumption of the leaves, buds, fruit, roots or consumption of plant sap. Aside from direct damage to the plants, sap sucking pests such as the Hemiptera also transmit economically damaging plant viruses (Gray,

2008). Currently the control of insect pests relies heavily on synthetic pesticides. Three quarters of pesticides sold act on one of five major targets in the insect nervous system and new compounds often act on the same targets (Casida, 2009). The limited modes of action for current pesticides combined with the injudicious use of pesticides has led to the rapid evolution of resistance to many of these compounds. The advent of genetic modification allowed for reduced reliance on synthetic pesticides by the introduction of insect resistance traits, such as crops expressing *Bacillus thuringiensis* (Bt) toxins. Bt expressing crops have provided significant increases in yield whilst reducing pesticide usage. However, they confer little protection to crops against attack by hemipteran pests and after twenty years of deployment around the globe issues such as resistance are beginning to reassert themselves.

This chapter focuses on exploring the current methods of insect control, aphids as a crop pest, V-ATPases as a novel target, the current work on pea albumin up to this point and the current state of GNA fusion protein technology. Current insect control methods are lacking due to a limited number of target sites and high selection pressure have led to increasing numbers resistant insect pests. Aphids are of particular interest here as they rapidly evolve resistance and the BT toxins mentioned above are largely ineffective. V-ATPase proton pumps may be a potential novel target, which up until recently have been largely ignored for use in insect control due to the lack of specificity of their chemical inhibitors. Recently pea albumin has been identified as a proteinaceous and specific inhibitor of insect V-ATPases. Finally, the enhancement of protein toxins through fusion to *Galanthus nivalis* agglutinin (GNA) has been demonstrated to enhance oral activity of small entomotoxins, and may be utilised to enhance the toxicity of pea albumin.

The hunt for novel and effective insecticidal molecules is an ever-ongoing arms race with the pest species. Novel compounds with novel target sites and modes of action are desperately needed. This thesis explores the potential for the use of pea albumin as a novel GNA fusion protein to do just that.

1.2 Current methods of insect control

1.2.1 Chemical control

The use of chemical insecticides on crops dates to the mid-19th century where inorganic compounds such as arsenic and boric acid were utilised for insect pest control. Advances in synthetic chemistry allowed the development of small molecule organic insecticides such as the organophosphates and pyrethroids (Le Goff and Giraudo, 2019). The majority of agricultural insect pest control is reliant heavily on synthetic small molecule insecticides with the quantity

of insecticides sprayed increasing 15 -20 fold between 1960 and 2004 (Oerke, 2005). The same article highlights how dependent modern farming is on these compounds with predicted yield losses in the absence of pesticides ranging from between 8.7 % in wheat to 36.8 % in cotton. Seventy five percent of the insecticidal compounds sold act on one of five target sites in the insect nervous system (Casida, 2009); namely voltage gated sodium (Na) channels, nicotinic acetylcholine receptors, acetylcholinesterase, glutamate receptors and ɣ-aminobutyric (GABA) receptors (Figure 1.1). For specific modes of action of different pesticide classes see Le Goff and Giraudo (2019).

Unfortunately, our reliance on these compounds has put intense selection pressure on insect pests and driven the rapid evolution of resistance. The first documented case of insecticide resistance was reported by Melander in 1914. The development of resistance by insect crop pests is facilitated by the lack of diversity in the target sites, coupled with indiscriminate use of insecticides (Bass *et al*., 2014; Feyereisen, 1995). More than 550 species of arthropod have been documented to have resistance to insecticides, with many having multiple resistances (Whalon *et al*., 2008). Examples of multiple resistances are insects such as *Plutella xylostella* (diamondback moth) and *Myzus persicae* (peach-potato aphid) that have respectively developed resistance to 92 and 76 unique active ingredients (Zhu *et al*., 2016).

Insecticide resistance is mainly grouped into four key mechanisms: Metabolic, target-site, penetration and behavioural resistance. The Insecticide Resistance Action Committee website (IRAC, 2019) defines these mechanisms as follows: Metabolic resistance: resistance through the detoxification, breakdown or sequestration of the toxic compound, resistant strains may possess higher levels or more efficient detoxification enzymes and this mechanism often has a broad spectrum of activity and is considered the most common form of resistance. Target-site resistance: resistance to toxic compounds through the mutation of the pesticides target-site such that it cannot interact, or the insecticidal effect is reduced. Penetration resistance: resistance which arises through reduced absorption of the toxic compound into the insect body through the development of barriers in the cuticle, this may protect from a wide range of insecticides and may be present in conjunction with other mechanisms. Behavioural resistance: The adaptation of insect behaviour to recognise and avoid the toxic compound this has been observed in several classes of insecticide.

An example of metabolic resistance is exists in *Nilaparvata lugens* (Brown plant hopper) in which the upregulation of a cytochrome P450 enhances resistance towards imidacloprid, a neonicotinoid. Carboxylesterase gene duplication in *M. persicae* is an example of target-site

resistance where increased expression of target-site allows normal function despite inhibition by organophosphates (Devonshire and Moores, 1982; Field *et al*., 1988). A further example of target site resistance is point mutations in the binding sites of insecticides can also confer resistance, in *Aphis gossypii* (cotton aphid) the mutation S431F in the acetylcholinesterase enzyme confers resistance to carbamates and organophosphates (Benting and Nauen, 2004). Other mutations have been observed which reduce the likelihood for voltage gated Na^+ channels to open and therefore reduces the efficacy of pesticides in the pyrethroid class (Martinez-Torres *et al*., 1999). Thickening of the cuticle to exclude deleterious compounds has also been observed in the field as a resistance mechanism that reduces the uptake and thus effectiveness of topically applied insecticides (Puinean *et al*., 2010; Strycharz *et al*., 2013).

Aside from issues of resistance, chemical insecticides can be hazardous to human health and the environment. For example, acute organophosphate poisoning can lead to vomiting, diarrhea, abdominal cramps, dizziness, eye pain, blurred vision, confusion, numbness, twitching, paralysis and death (Koureas *et al*., 2012). Chronic exposure to organophosphates has been suggested to increase the risk of developing Parkinson's disease. However, the data are inconclusive, and more work is required (Freire and Koifman, 2012). Furthermore, increased risk of developing numerous cancers including leukemia, lung ovarian and colon cancer, have been associated with increased pesticide exposure (Sabarwal *et al*., 2018). Pesticide exposure may also result in reproductive abnormality in mammals. In mice, exposure to carbamate pesticide, Carbosulfan, has resulted in increased sperm abnormalities, and supplying Cypermethrin, a pyrethroid, in the drinking water of rats resulted in significant decreases in fertility (Elbetieha *et al*., 2001; Giri *et al*., 2002). Annually, over 1 million human deaths and incidents are attributed to insecticide poisoning. Developing countries are disproportionally affected (Aktar *et al*., 2009).

Unsurprisingly, the use of broad-spectrum insecticides has negative effects on biodiversity across Europe (Geiger *et al*., 2010). Documented effects on beneficial species are not just limited to mortality but can range from the reduced ability to capture prey, reductions in oviposition, along with developmental and reproductive impairments (Ndakidemi *et al*., 2016). Further to this, negative effects upon pollinators are increasingly recognised. For example neonicotinoid pesticides have been shown to have deleterious effects upon bee health and memory (Woodcock *et al*., 2017). The mass spraying of insecticides has also had an impact on vertebrate species: neonicotinoid pollution has been linked to declines in bird populations (Goulson, 2014). Recently, sublethal doses of the neonicotinoid imidacloprid when ingested by white-crowned sparrows during migration have been reported to result in reductions in food

consumption, mass and fat content. Additionally, departure from the feeding site was delayed, which can affect survival and reproduction. The effects of pesticide application are not limited to the site of application, the pollution can spread through run off and enter streams and rivers and groundwater where at least 140 different pesticides have been found (Aktar *et al*., 2009). Ippolito *et al*., (2015) predict that water bodies in 40 % of the earth surface have a high risk of adverse effects due to insecticide contamination. Negative side effects such as these have led to the retraction or banning of multiple effective insecticides, such as the neonicotinoids clothianidin, imidacloprid and thiamethoxam in field crops and the complete banning of dichlorodiphenyltrichloroethane (DDT) (Abdourahime *et al*., 2019; Beard, 2006; European Commission, 2018a, 2018b, 2018c).

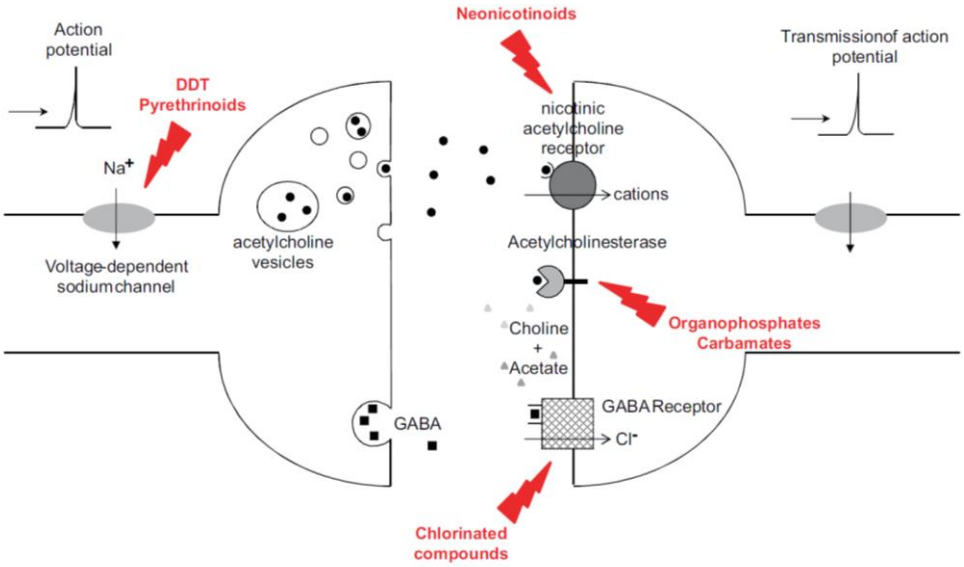


Figure 1.1 Targets sites of action of major insecticide classes at a synapse. Image source: Le Goff and Giraudo (2019).

1.2.2 Biological control

Synthetic chemicals, although the principal method of crop protection, are not the only way in which pest insects can be controlled. Biological control is defined as the use of natural enemies to reduce the damage caused by pest species (DeBach and Rosen, 1991). There are three main but overlapping categories in biological control: classical, augmentative and conservation biological control. The categories are designated based on how natural enemies are utilised. In classical biological control the importation and establishment of non-native natural enemies is utilised to control an indigenous or non-native pest population. Success utilising this method has been documented as far back as 1889, where *Rodolia cardinalis*, the Australian vedalia lady beetle, was introduced into Californian orange groves to control *Icerya purchasi* (cottony cushion scale). This is regarded as the first instance of modern biological control (Howarth, 1991). The second category of biological control, augmentative control relates to the augmentation of previously established species, be they native or exotic, through the periodic release of predator species (van Lenteren, 2012). Periodic release of *Trichogramma* prarsitoid wasp eggs has been used to great effect to control various lepidopteran pests including *Ostrinia nubilalis*, the European corn borer, and *Spodoptera frugiperda,* fall armyworm (de Lourdes Corrêa Figueiredo *et al*., 2015; Smith, 1996). Augmentative control is often complimented by conservation biological control, where the environment is manipulated to be conducive to the maintenance and survival of the natural enemies which have been introduced or are already present. However conservation biological control can be effective on its own as well as in combination. Manandhar and Wright, (2016) found that the interplanting of sweet corn with cow pea significantly increased *Trichogramma* parasitism of *Helicoverpa zea*, corn earworm, eggs and the number of thrips were greatly reduced.

As of 2012 there were at least 230 species of natural enemy are available for use on a commercial scale (van Lenteren, 2012). This practice has been shown to be successful in some areas of agriculture, including fruit orchards, maize, cotton, soybean, vineyards and glasshouses (van Lenteren and Bueno, 2003). Throughout Europe the uptake of biological control through the release of natural enemies in glasshouses is increasing, with lacewing, hoverfly and ladybird all being utilised for the control of aphids (Pilkington *et al*., 2010). Successful examples of biological control include the introduction of the parasitic wasp *Encarsia formosa* into glasshouses to control the greenhouse whitefly (*Trialeurodes vaporariorum)*. This wasp was used as a control method as early as the 1920s but fell out of favour for chemical pesticides after the second world war. When the issue of chemical resistance exerted itself in the 1970s the use of

E. formosa found a renewed interest and is currently used effectively throughout Europe as an augmentative biological control (van Lenteren *et al*., 1996).

However, the artificial introduction of natural enemies must be done carefully. Disregard has resulted in the introduction of invasive species such as the Harlequin ladybird (*Harmonia axyrids*) which has had significant impacts on native ladybird populations and insect predator species in the United States of America (USA) (CABI, 2019; Mizell III, 2007). Further in parts of the USA, *H. axyridis* is now considered a pest of orchard fruit and a nuisance to humans (Koch, 2003). A further and perhaps more famous example of biological control backfiring was the introduction of the cane toad, *Rhinella marina*, into Queensland, Australia in 1935. The intention was to control the cane beetle, this however failed due to spatial separation of their habitats, beetles on the stems and the toads on the ground, the cane toads instead resorted to eating other flora and fauna. The arrival of the cane toad has also resulted in reductions in predator species as they are fatally toxic to many taxa (Jolly *et al*., 2015). Improper due care and diligence can lead to devastating results as noted above. As a result of these negative impacts regulations have been put in place , such as the 'Guidelines for the export, shipment, import and release of biological control agents and other beneficial organisms' to reduce and prevent these negative impacts from occurring.

1.2.2.1 *Bacillus thuringiensis*

A notable example of successful biological control is the spraying of *Bacillus thuringiensis* to control insect pests. *Bacillus thuringiensis* (Bt) is a soil-dwelling gram-positive bacterium. It was initially identified in 1901 by Ishiwatari whilst investigating wilt disease in silk worms. In the 1920s the insecticidal activity of the crystals produced by Bt were discovered and successful field trials utilising Bt spores to control the European corn borer in the 1930s (*Ostrinia nubilais)* facilitated commercialisation of the first Bt bioinsecticide Sporeine which was brought onto the market in 1938 (Sanahuja *et al*., 2011). *Bacillus thuringiensis* produces two insecticidal pore forming toxins during its stationary growth phase, the Crystal (Cry) or Cytolytic (Cyt) toxins (Bravo *et al*., 2011). More than 500 different *cry* gene sequences have been discovered and placed into 67 groups (cry1-67) based on their primary amino acid sequence, each individual toxin is generally only toxic within one order (Crickmore *et al*., 2014). These toxins are highly toxic to insects in the orders of Lepidoptera, Coleoptera and Diptera.

The Cry toxins are produced as a three domain proproteins and secreted as a crystalline water-soluble compound. Their mode of action has been studied mostly in the context of

lepidopteran insects (for review see Bravo *et al*., 2007). In order for the toxin to function the N-terminal region, 25-58 amino acid residues, must be cleaved by insect midgut proteases. The cleavage results in a 60-70 kDa protease-resistant protein (Bravo *et al*., 2007). The removal of the N-terminal propeptide may provide a level of specificity as lepidopteran insects have an alkali gut environment which favours serine proteases when compared to coleopterans which have an acidic gut environment which likely favours cysteine proteases (Dow, 1992). The requirement for cleavage may account for part of Bt toxins reduced activity towards hemipteran pests which generally have low proteolytic activity, preactivation with trypsin conferred some activity towards aphids (Porcar *et al*., 2009). After the cleavage of the N-terminus has occurred the activated toxin binds, via domain II, to a receptor in the insect gut epithelial brush border membrane. There are a number of proteins which have been shown to act as receptors for Bt cry toxins which include cadherin-like proteins (CADR), an aminopeptidase-N and an alkaline phosphatase, all of which are anchored to the membrane by a glycosylphosphatidyl-inositol (Jurat-Fuentes and Adang, 2004; Knight *et al*., 1994; Vadlamudi *et al*., 1995). Binding to specific proteins is widely suspected to be the major determinant for the specificity observed in the Cry toxins (Pigott and Ellar, 2007). Once bound to these membrane proteins the toxin undergoes another N-terminal cleavage to form a pre-pore oligomeric structure (Gómez *et al*., 2002). Subsequently domain I is inserted into the membrane forming a pore, leading to cell swelling and lysis, and eventually death of the insect (Bravo *et al*., 2007; de Maagd, 2001). Identification of similar binding proteins in other insect orders and the conserved three domain fold present in all Cry toxins, has led to the hypothesis that the mode of action of Cry toxins is conserved across orders (Bravo *et al*., 2011). The Cyt toxins function similarly but do not require a protein binding site and interact directly with the membrane lipids rather than a protein receptor (Promdonkoy and Ellar, 2003; Thomas and Ellar, 1983).

1.2.3 Insect resistant transgenic crops expressing toxins from *Bacillus thuringiensis*

Despite Bt based insecticide sprays being developed and available in the 1930s, Bt toxins were not widely used outside of organic farming until the advent of genetic modification. The development of transgenic crops expressing Bt Cry toxins has been a powerful tool in pest control. The first Bt potato expressing Cry1ac was commercialised in 1996 by Monsanto (Castagnola and Jurat-Fuentes, 2012). Transgenic cotton expressing Cry2Ab has been shown to confer resistance to numerous pests including the cotton bollworm, tobacco budworm and pink bollworm (Roh *et al*., 2007). Alongside increased resistance to crop pests, Bt crops have also been associated with better grain quality in Bt maize as compared to crops produced using

traditional farming techniques (Hellmich and Hellmich, 2012). After the commercialisation of Bt cotton in China in 1997 there has been a steady increase in the adoption of the transgenic crop such that by 2008, six of the eight major cotton growing regions in China were growing only Bt cotton. Further, the money spent on pesticides has decreased by 35% alongside decreases in labour required to farm this land (Qiao and Yao, 2015). A similar trend can be seen in India where in 2010 it was reported that 86 % of all cotton grown (covering an area of 10 million hectares) was transgenic Bt cotton, (James, 2010). As of 2017, cultivated crops expressing genes from *B. thuringiensis* were grown on 101 million hectares of farmland globally (ISAAA, 2017). The success of Bt crops is undeniable.

Not only are there economic benefits gained from reduced pesticide usage, there are also health and environmental benefits as there are less potentially hazardous compounds being sprayed into the environment and farmers are less exposed to the negative effects of insecticides (Purcell and Perlack, 2004). Further there are yet to be any studies which demonstrate that Bt crops have any significant negative effects upon larval or adult stages of the beneficial pollinator *Apis mellifera* (Duan *et al*., 2008; Hendriksma *et al*., 2011).

Although there are a considerable number of benefits to the use of Bt crops there are large concerns about the development of insect resistance toward these toxins, especially as many of the Bt toxins have been in use since the 1920s (Lemaux, 2008). The *Plutella xylostella* (diamondback moth) has been shown to develop resistance to Bt in the field (Tabashnik, 1994). The Asian corn borer has been shown to be able to develop 100-fold resistance to Cry1Ab, over only 35 generations, and after 51 generations the insects could survive on Cry1Ab expressing maize. Not only did this confer resistance to Cry1Ab the insects had gained cross-resistance to Cry1Ah and Cry1Ac at levels of 131-fold and 36-fold, respectively (Xu *et al*., 2010). Cross resistance to toxins which the insects have never been exposed could present potential problems for the use of Bt moving forward. However, Xu *et al*., (2010) does state Cry1Ab and Cry1Ah share 88 % sequence similarity and this may be the cause for the observed cross resistance. Another worrying example of cross resistance in *P. xylostella*, where a single recessive gene confers resistance to four Bt toxins (Tabashnik *et al*., 1997). The second generation of Bt crops has utilised gene pyramiding (i.e. the expression of multiple resistance traits), to delay resistance development, with Monsanto's Bollgaurd II expressing both Cry1Ac and Cry2Ab toxins. The deployment of gene pyramided crops expressing two or more toxins is not a silver bullet for resistance as research has demonstrated that prior exposure to single toxin expressing crops can substantially reduce the time required for resistance to develop to the dual

trait crop (Zhao *et al*., 2005). Application of correct farming practices could reduce the speed at which resistance develops through the use of refuges, where planting areas of non-modified crops or plants act to reduce the selection pressure on target pests for the expression of resistance genes (Gryspeirt and Grégoire, 2012). In India where the use of non-Bt cotton refuges are scarce, the pink bollworm has shown to be resistant to cotton that expresses both Cry1Ac and Cry2Ab toxins. In the USA and China, where the refuges are more common this resistance is much less prevalent (Tabashnik and Carrière, 2019).

Aside from concerns of resistance, Bt crops have one further major drawback in that the Bt toxins are only marginally active toward hemipteran pests (Chougule and Bonning, 2012). The reduction in the use of pesticides has in some cases actually increased populations and damage caused by hemipteran pests on Bt expressing crops (Men *et al*., 2005). As *B. thuringeinsis* exists in the soil and is splashed onto foliage by rainfall, hemipteran pests would not naturally be exposed to the Bt toxins as sap suckers they pierce the exterior of the plant rather than consuming it, thus there would be no selective pressure for *B. thuringiensis* to evolve toxins which target hemipteran insects (Schnepf *et al*., 1998). Further hypotheses suggest perhaps a lack of binding sites or non-optimal conditions for protein activation may exist in the hemipteran gut (Chougule and Bonning, 2012; Lawo *et al*., 2009; Li *et al*., 2011; Porcar *et al*., 2009). In support of this hypothesis Chougule *et al*., (2013) managed to retarget Cyt2A toward pea and peach potato aphids by the addition of a protein domain which will bind to aminopeptidases in the aphid gut. Similarly the introduction of cathepsin L and B sites into Cry4a between domains I and II, also resulted in enhanced toxicity towards *A. pisum* (Rausch *et al*., 2016).

1.2.4 Integrated pest management

Integrated pest management (IPM) is defined by the FAO (2019) as “the careful consideration of all available pest control techniques and subsequent integration of appropriate measures that discourage the development of pest populations and keep pesticides and other interventions to levels that are economically justified and reduce or minimize risks to human health and the environment.” This is generally achieved through the combination of crop protection techniques such as pest monitoring, biological control, mechanical separation, modification of cultural practices, the use of resistant varieties and as a final resort, pesticides. Pesticides are only applied when the monitoring indicates that the pests may be reaching a predefined threshold.

The European Commission (2017) defines eight principles that are the cornerstone of the IPM approach. The first is that prevention and suppression of harmful organisms should be

achieved or supported at all times by non-chemical means such as crop rotation, physical barriers, prevention of spread on machinery, the use of resistant cultivars where appropriate and the protection and enhancement of beneficial organisms amongst other things. Second, harmful organisms must be identified and monitored utilising adequate tools, such as field observation and scientifically sound warning, forecasting or early diagnosis systems in combination with advice from qualified advisors. Third, based on the evidence gathered through monitoring the user must decide whether to apply protection measures. therefore, scientifically determined threshold values are key in the decision-making process and these must consider the pest in question, the region, crops, and climate conditions. The fourth principle states that biological, physical and other non-chemical methods are to be used preferentially to chemical control if they provide effective pest control. Fifth if pesticides are to be applied, they must be as specific as possible for the target pest and have the least side effects in regard to the environment, human health and non-target organisms. Sixth, application of pesticides and other interventions should be kept to the lowest necessary level. Seventh, if there is a risk of resistance against a protection measure and where the level of pest organisms requires repeated application of pesticides anti-resistance strategies should be employed including the use of multiple pesticides with differing modes of action. Finally, and perhaps most importantly, is based on the records of use and monitoring the success of the protection measures applied should be evaluated and recorded for future reference.

A key goal of IPM is the reduction of pesticide usage whilst maintaining crop growth at an economically viable level. A good example of this combinatorial approach to pest management is reported by Lu *et al*. (2009). In this study *Apolygus lucorum* a hemipteran pest was maintained below an economically damaging threshold whilst reducing insecticide usage to below 70 % of conventional farming. This was done by introducing a plant on which *A. lucorum* would feed preferentially as 10 % of the crop area, the other 90 % was Bt cotton and therefore likely resistant to coleopteran and lepidopteran pests. Insecticide was not sprayed on the cotton and only applied on the mung beans once *A. lucorum* reached a predefined threshold. The application of IPM in the citrus orchards of Australia, where the adoption of classical biological control has reduced pesticide usage by up to 75 % whilst reducing the costs of management (Zalucki *et al*., 2009). Further success utilising an IPM approach is reported in Java where pesticide usage is reduced through more targeted application of pesticide usage in rice paddies (Mariyono, 2008). However, there is still some debate on whether the drop in pesticide usage seen since the introduction of IPM in the latter half of the 20th century is due to IPM or due to

the introduction of genetically modified crops and low volume pesticides (Peshin and Zhang, 2014). Despite the debate IPM is currently the driving approach for the development of new agricultural practices. The European Union, American Environmental Protection Agency and the National Farmers Union websites all recommend the use of IPM strategies (European Commission, 2017, NFU 2017; US EPA, 2014). As such it is important to consider and demonstrate that novel control methods are compatible with this approach. Novel and specific biopesticides, such as fusion proteins, could and do play a role in IPM through utilisation as alternatives to spraying synthetic insecticides. Plants could be sprayed with them or genetically modified to express them. In line with the principles of IPM, this would only be the case where biological, physical and non-chemical methods were deemed ineffective. They may be of particular relevance in aligning to the IPM principles due to their potential for use in anti-resistance strategies that require multiple pesticides with differing modes of action, and in cases where they have higher pest specificity than synthetic pesticides.

1.3 Aphids

Aphids are small sap sucking insects of the order Hemiptera or 'True bugs'. The Hemiptera are one of the largest orders of insects comprised of 90,000 species and can range in size from 1 mm up to 15 cm, and includes cicadas, leafhoppers, planthoppers, aphids and stink bugs. The defining characteristic of the Hemiptera is that the mouthparts, the mandibles and the maxillae have formed into a dual channel stylet where one channel delivers saliva and the other uptakes fluids. Aphids fit into the suborder of Sternorrhyncha, which are generally characterised by their possession of a particular structure of filter chamber and a rostrum that arises between the bases of their front legs (Cranston and Gullan, 2009). Aphids have a huge range of diversity in their host specificity, with certain species practicing monophagy, some species preferring certain sub-groups of crops such as the pea aphid *Acrythosiphon pisum* feeding on legumes with a preference for *Vicia faba*, and some species are polyphagous, for example *Myzus persicae* (van Emden *et al*., 1969; McVean and Dixon, 2002).

Aphids generally feed passively on the vascular tissue of plants, particularly the phloem sap. The high pressure within the sieve elements drives the liquid into the aphid stylet which is elongated and can pierce plant tissue to reach the vasculature. Aphid feeding is aided by the production of two different types of saliva known as gelling and watery (Tjallingii, 2006). The former is jelly-like and proteinaceous and coats the stylet providing a physical barrier between the stylet and the plant tissue (Miles, 1999). The second, watery saliva, is injected into the phloem and contains lytic enzymes and proteases to aid in the digestion of nutrients present in

the phloem (van Bel and Will, 2016; Furch *et al*., 2015). It also contains effector proteins which can attempt to circumvent plant defences and prevent the wounding response (Will *et al*., 2007). For example the Mp55 protein secreted in the saliva of *M. persicae* reduces callose deposition in the phloem of plants under attack by aphids (Elzinga *et al*., 2014). This feeding lifestyle causes little physical damage to the plant as compared to chewing pests, and often helps to avoid plant defences, some chemical pesticides, and potentially reduces their exposure to transgenic toxins.

Feeding on phloem sap is an unbalanced diet which is exceedingly high in sucrose and low in nitrogen, the majority of which is in the form of free amino acids. The high quantity of small sugars, such as sucrose in the phloem sap leads to high osmotic pressure in the gut of the aphids which can be lethal, therefore dietary sugars exceeding dietary needs are converted into oligosaccharides which are then excreted in the honeydew (Douglas *et al*., 2006). A further constraint on aphid growth is the low level of essential amino acids present in phloem sap, for example in *V. faba* phloem sap only the non-essential asparagine is in excess of what is required by *A. pisum* (Douglas, 2006). Aphids are therefore reliant on endosymbionts to produce their essential amino acids (Hansen and Moran, 2011). For example, *A. pisum* contains some of the genes responsible for key reactions in the synthesis of certain amino acids which are not present in its symbiont *Buchnera aphidicola*, but is reliant on *B. aphidoicola* for the remainder of the pathway (Wilson *et al*., 2010). The aphid gut is a finely balanced environment reliant on its symbionts for correct regulation of osmotic pressure and disruption of this system may provide an ideal target for novel pesticide development.

Aphid lifecycles can vary tremendously between species but generally, follow one of two major types. Heteroecious (host-alternating) and monoecious (non-host alternating). The former involves having a minimum of two plant host species, the species on which they feed is determined by season; the primary host is where eggs are laid and where aphids live during the winter, and then migrate to the secondary host for the summer months, where several parthenogenic generations take place. Both of these two life cycles can either be entirely parthenogenic, reproducing asexually, or they may reproduce parthenogenically most of the year and then form a sexual morph and reproduce sexually laying over wintering eggs. Aphids which break their cycle of parthenogenic reproduction to have a sexual stage are termed holocyclic. Non-host alternating species live on the same host year round (van Emden and Harrington, 2007). Often in regions where the primary host is absent holocyclic species can survive using an anholocyclic life style, a good example of this is *M. persicae* which usually undergoes its sexual stage on peach trees (*Prunus persicae)*, however as they are uncommon in

the United Kingdom, *M. persicae* rarely completes the full holocyclic life cycle and overwinters on one of its many secondary hosts as parthenogenic females (Williams *et al*., 2000).

Despite only approximately 100 species of aphids feeding on crops of agricultural importance they can cause significant economic damage (Blackman and Eastop, 2007). Damage is incurred in two ways. The most important and devastating way in which aphids cause significant economic damage is through the transmission of plant viruses with an estimated 275 of 600 insect vectored plant viruses being transmitted by aphids, although this estimate may be outdated (Nault, 1997). Within the 275 viruses several are of significant economic importance such as Barley yellow dwarf disease (BYDD), Potato virus Y (PVY), Turnip mosaic virus (TuMC), Potato leaf roll virus (PLRV) (D'arcy, 1995; Tomlinson, 1987). Aphids as phloem feeders also cause direct damage to the plant by diverting nutrients from the plant to fuel their own growth and reproduction, and while each individual aphid is small, their parthenogenic lifestyle means that populations can rapidly reach economically damaging levels. For example, the doubling time of the grain aphid, *Sitobion avenae* under ideal conditions can be as low as 55 h, with a population of 100 *S. avenae* able to reach 3.2 million in just over a month (Simon *et al*., 1991). Aphids are also capable of producing winged morphs in response to population size, host quality and the presence of predators, in order to disperse to find more optimal conditions (Mehrparvar *et al*., 2013). This ability to rapidly reproduce and . Dedryver *et al*., (2010) presents statistics suggesting that in Europe direct damage by aphids is responsible for mean annual losses of 700,000 tons (t) of wheat, 850,000 t of potatoes and 2,000,000 t of sugar beet.

This thesis focuses on three model species, with varying lifestyles and varying economic importance, but all three species are listed in Emden and Harrington (2007) as being in the top 14 economically important aphid species. The three aphid species used in this project are *Acrythosiphon pisum* (Pea aphid), *Myzus persicae* (Peach-potato aphid) and *Sitobion avenae* (grain aphid).

1.3.1 *Myzus persicae*

Myzus persicae or the peach-potato aphid is distributed worldwide and considered one of the most important hemipteran agricultural pests, with a secondary host range of over 400 species across 40 different plant families (Bass *et al*., 2014; CABI, 2018a). The peach-potato aphid displays a heteroecious lifecycle with its sexual stages taking place specifically on *Prunus persicae* (peach trees). This pest has been shown to transmit over 100 different plant viruses many of which are agriculturally important such as PLVR and BYDD (Blackman and Eastop, 2007; Ponsen, 1972). This pest has a remarkable ability to evolve resistance to chemical pesticides

through various mechanisms, particularly overproduction of detoxification enzymes, and mutations in the binding site of insecticidal molecules (Bass *et al*., 2011; Devonshire and Moores, 1982). This study used *M. persicae* in the assessment of the oral toxicity of fusion proteins towards a model generalist species.

1.3.2 *Acrythosiphon pisum*

Acrythosiphon pisum (Pea aphid) is a relatively large aphid pest often reaching 4mm across. It is a legume specific pest and its host plants are mainly legumes of the *Fabaceae* tribes Genistae, Trifoliae and Fabeae. It is generally non-host swapping and holocyclic with both the sexual and asexual stages occur on the same plant species (Blackman and Eastop, 2007). It has been detected on legume crops worldwide (CABI, 2018b) and vectors up to 30 viruses including two economically important viruses of legume crops, namely Pea enation mosaic virus (PEMV) and Bean leaf roll virus (BLRV) (Hodge and Powell, 2010; Schwinghamer *et al*., 2009). This insect is often used as a model aphid species in laboratory studies. The genome of *A. pisum* contains a lower diversity of detoxification genes as compared to the generalist *M. persicae* (Ramsey *et al*., 2010). This decrease in detoxification proteins is believed to be due to the reduced quantity of plant defence compounds a specialist species would encounter. In this study *A. pisum* was used as a model specialist species when testing the oral toxicity of fusion proteins.

1.3.3 *Sitobion avenae*

Sitobion avenae (grain aphid) is a semi-specialist species which colonizes numerous species of Poaceae, including all cereal crops grown in temperate climates and is a particularly important pest of wheat in western Europe (Larsson, 2005), but has been identified in all continents except Australia and Antarctica (CABI, 2018c). This pest is a particularly effective vector of BYDV. This species does not display host alteration as displayed by *M. persicae* and both its sexual and non-sexual stages occur on its primary host. In Europe and other cooler climates *S. avenae* produces sexual morphs but in warmer climates such as India it is parthenogenic year-round (Blackman and Eastop, 2007).

1.4 V-ATPase proton pumps

The vacuolar adenosine triphosphatase (V-ATPase), is a highly conserved large multi-subunit membrane associated holoenzyme that functions as a rotary proton pump to move protons up a concentration gradient utilising the energy released by the hydrolysis of ATP (Nelson *et al*., 2000). It is generally considered the polar opposite of the F-ATPase which utilises

a proton gradient to facilitate the formation of ATP by catalysing the phosphorylation of ADP, however they are evolutionarily related (Beyenbach and Wieczorek, 2006; Nishi and Forgac, 2002). V-ATPases are ubiquitous throughout eukaryotic membranes and are responsible for many secondary transport functions alongside the correct functioning of organelles such as the Golgi apparatus (Forgac, 2007).

1.4.1 V-ATPase structure and mechanism

V-ATPases being large holoenzymes (depicted in Figure 1.2) are made up of 14 different polypeptides which form two complexes, V_1 and V_0. V_1 is soluble and located on the cytoplasmic side of the cell membrane and made up of subunits $A_3B_3CDEFG_2H_{1-2}$ with a mass of 400-600 kDa. The second complex is membrane associated and formed of several transmembrane subunits, $ac_4c'c''de$, with a mass of 150-350 kDa. The subunits of V_1 and V_0 are distinguished by uppercase and lowercase letters. The cytoplasmic V_1 subunit is responsible for the hydrolysis of ATP into ADP and inorganic phosphate. The energy released is used to drive the rotor of the second complex V_0 which facilitates proton transport across the membrane. The V_1 complex consists of three distinct structures, the first is a globular structure of three sets alternating A and B subunits forming a ring, each of which mediates ATP hydrolysis. A central rotating stalk made up of D and F and finally a peripheral stalk, made up of B and C, both of which connect V_1 to the central c-ring structure of V_0. The V_0 subunit mostly consists of a hydrophobic-ring made up of six or more c-subunits, which rotate within the a and e-subunits. The widely accepted functional model suggests that V-ATPases are made up of a rotor (mobile) and a stator (stationary). Subunits c, D and F make up the rotor and everything else is considered the stator (reviewed by Beyenbach and Wieczorek, 2006).

Proton transport across the membrane relies on the structure of the V_0 complex which contains two half proton channels formed by the close proximity of subunits a and c, one exposed to the cytosol and one to the extracellular fluid (Grabe *et al*., 2000). Each subunit of the c-ring also contains a proton binding glutamate residue, clockwise rotation of the c-ring, driven by conformation changes of subunits A and B due to ATP hydrolysis, allows transport of the proton from the cytosolic channel to the extracellular channel (Meier, 2005; Murata, 2005). Movement of the protons from the cytosol to the intracellular compartment results in acidification of the intracellular compartments, and plays a key role in the maintenance of gut pH and secondary transport in insects (Klein, 1992; Marshansky and Futai, 2008; O'Donnell, 2017).

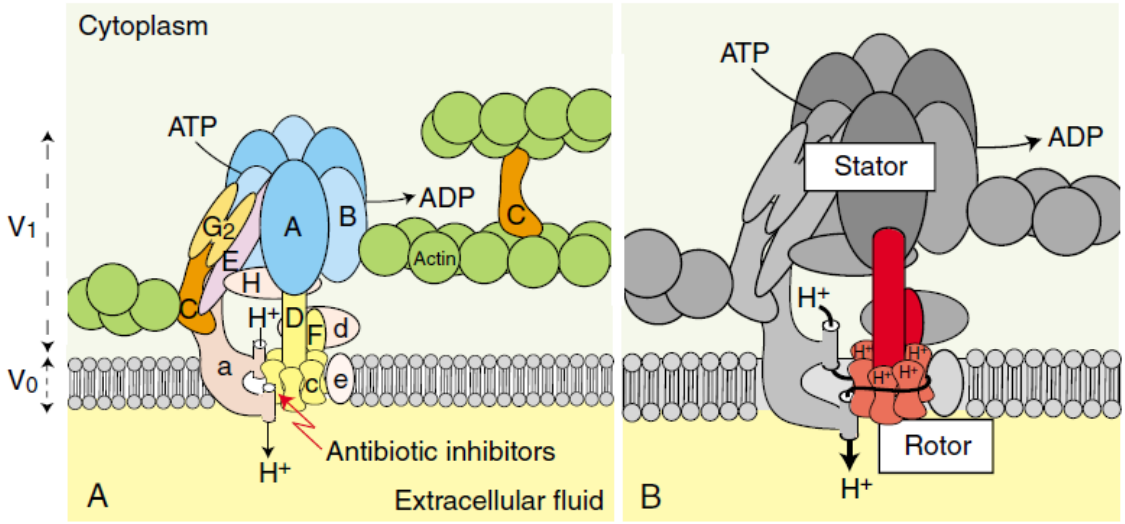


Figure 1.2 Model of the eukaryotic V-ATPase proton pump expressed in a cell membrane. (A) Molecular model. The cytosolic V_1 complex consists of eight different subunits identified with uppercase letters (A-H), G exists as a dimer G_2. The integral membrane bound V_0 complex is made up of four different subunits identified with lowercase letters (a,c,d,e). Subunit c and its isoforms form a hydrophobic proton binding ring. (B) Mechanistic model. V_0 is connected to V_1 by a central rotating shaft (D and F) and a peripheral stationary shaft (C, E, G, H and a). The rotor consists of the central rotating shaft and the c-ring of V_0 (red) the remaining subunits make up the stator (grey). Hydrolysis of ATP drives rotation of the rotor. Subunit a provides two half proton channels which are offset, rotation of the c-ring moves protons from one channel to the other (Beyenbach and Wieczorek, 2006).

1.4.2 V-ATPases as a target site for insect control

In recent years many studies have investigated targeting the V-ATPase proton pumps in the insect gut epithelium as a method of insect control. Most of these studies have investigated using RNA interference (RNAi) to knock down the expression of a key subunit of these proteins. Results have been variable, but where the RNAi has been shown to function results are promising. The A subunit in V_1 has been a popular and effective target. Oral delivery of dsRNA encoding for subunit V_1 to western corn rootworm, small hive beetle and whitefly showed significant decreases in survival and development as compared to controls (Baum *et al*. 2007; Upadhyay *et al*. 2011; Powell *et al*. 2017). RNAi mediated knock down of the e-subunit in Colorado potato beetle, *Leptinotarsa decemlineata* also resulted in decreased larval survival and development (Fu *et al*., 2014).

Bafilomycin, a specific inhibitor of V-ATPase pumps, was patented in the USA for use as a pesticide, where it was stated that concentrations of 0.004 % (v/v) sprayed onto cabbage plants caused 100 % mortality of larvae of Diamond backmoth, *Plutella maculipennis,* after 14 days (Bowman *et al*., 1988; Hagenmaier *et al*., 1985). The above is a small subsection of literature that exists on the detrimental effects of knocking down subunits of insect V-ATPases, the success of these studies and the toxic effect of inhibition of V-ATPases suggest it may be a good target for further development. An insecticidal protein from pea seeds known as PA1b which has been shown to specifically inhibit insect V-ATPases, is discussed further in the next section and is a promising candidate for further development (Chouabe *et al*., 2011).

1.5 Pea albumin

1.5.1 Initial discovery and characterisation and early work

A low molecular weight sulphur rich pea albumin protein was discovered in *Pisum sativum* and its amino acid sequence was elucidated in 1985, its initial function was thought to be as a novel sulphur storage protein, although it did not show homology to existing storage proteins (Gatehouse *et al*., 1985). The following year the gene sequence and expression pattern of the pea albumin protein (PA1) during development were elucidated by Higgins *et al*., (1986). Higgins *et al*., (1986) also demonstrated that the PA1 protein is expressed in seeds as a preproprotein (11 kDa); the signal peptide is cleaved in the endoplasmic reticulum and then the protein is subsequently directed to a vacuole-like protein storage organelle, where it is cleaved into its mature proteins, PA1a (6 kDa) and PA1b (4 kDa). These two proteins contain up to 50 % of the total sulphur in the seeds of *P. sativum* and are generally regulated at the transcriptional

level. Sulphur availability has been shown to be a key factor of the regulation of these genes, with sulphur deficient plants showing reduced quantities of PA1 mRNAs (Morton *et al*., 1993).

Further work and investigation into the gene regulation and properties of the pea albumin proteins was conducted through the late 1980s and early 1990s but no reports of insecticidal activity appear until 1993 when screens of four species of cereal weevil, *Sitophilus zeamais, S. oryzae, S. granaries* and *S. linearis* were fed on split peas. Of the four species tested only the *S. oryzae* strain China showed any ability to feed or develop on split peas, the other three species all reported 80 – 100 % mortality within 40 days (Delobel and Grenier, 1993). This toxicity was eventually attributed to the PA1b peptide and its use as an insecticidal toxin was patented in 1998 (Delobel *et al*., 1998). The PA1 proteins were also under investigation for use as nutritional enhancers that would be stable in the rumen of grazing animals. The first attempt to produce PA1 in a heterologous expression system was an attempt to use PA1 as a nutritional enhancer which would be stable in the rumen of grazing animals. and the authors demonstrated that PA1 is expressed and processed correctly in *Trifolium repens* (white clover), however despite high quantities of mRNA being found in leaf tissue very little protein was detected (Ealing *et al*., 1994).

1.5.2 Biological activity of PA1b

After the initial discovery of the insecticidal activity of PA1b against *Sitophilus* spp. very few papers have examined the range of insecticidal activity of PA1b. The initial patent of PA1b displays figures which indicate toxicity toward *Acrythosiphon pisum* (pea aphid) and *Ephestia kuehniella* (Mediterranean flour moth) (Delobel *et al*., 1998). Gressent *et al*., (2007) went on to test the range of toxicity of PA1b against a series of different species from different insect orders. The authors tested three different species of aphid, specifically *A. pisum, Aphis gossypii* (cotton aphid) and *M. persicae* (green peach-potato aphid). PA1b was very effective against *A. pisum* causing 100 % mortality after 6 days when fed at 0.5 mg mL^{-1} in artificial diet, but less effective against *A. gossypii* with only 30 % mortality at 1 mg mL^{-1} recorded and finally *M. persicae* was insensitive with no mortality observed at concentrations up to and including 1 mg mL^{-1}.

In contrast to the information provided in the PA1b patent, the survival and development of the lepidopteran pest *Mamestra brassica* (cabbage moth) was not affected when larvae were fed PA1b at doses up to 1000 mg g^{-1}. Personal communication referenced in Gressent *et al*., (2011) also suggests that other lepidopterans such as *Spodotera littoralis* and *Ostrinia nubilalis* are insensitive to PA1b. Interestingly despite the claim that *S. littoralis* are insensitive, cultured cells from this insect (SF9 cells) displayed sensitivity to low doses of PA1b

(Rahioui *et al*., 2014). The dipteran pest and vector of West Nile virus and Japanese encephalitis *Culex pipens* larvae were rapidly killed by concentrations of 250 µg mL^{-1} of PA1b added to water in which they were kept. Again, personal communication in Gressent *et al* (2011) suggests that PA1b may also be toxic to the yellow fever mosquito *Aedes aegypti*, however no published data are available.

Gressent *et al*., (2007) also tested PA1b against numerous coleopteran species, 3 strains of *S. orazyae* showed LC_{50} values of approximately 500 µg g^{-1}, *S. zeamais* showed an LC_{50} of 234 µg g^{-1}. The coleopteran flour beetle *Tribolium castaneum* survived for 15 days on diets containing 1000 µg g^{-1} of PA1b, conversely the ladybeetle *Harmonia axyridis* was particularly sensitive, with relatively low (250 µg g^{-1}) concentrations of PA1b found to cause major decreases in survival and disrupted development, with only 20 % of insects surviving to pupation and of those only 50 % pupated as compared to 80% control survival and 100 % pupation. By contrast, parasitic wasps of the genus *Trichogramma* were also reported to be unaffected by PA1b.

Gressent *et al*., (2007) also tested this toxin on the bacterial species *Pseudomonas aeruginosa, Staphylococcus aureus, Enterococcus faecium* and *Escherichia coli* and the fungi *Candida albicans, C. glabrata* and *Aspergillus fumigatus*. No effect was detected on any species up to 48 hours after the addition of the toxin to the respective culture medium at 64 µg mL^{-1}. PA1b also showed no activity when exposed to mammalian cells (Rahioui *et al*., 2014).

1.5.3 Mode of action and the structure of PA1b

The mode of action of PA1b has been unknown until recently. A high affinity binding site was discovered in the gut membrane of susceptible *Sitophilus spps,* four strains of *S. orazyae* which were previously shown to have a monogenic resistance to PA1b displayed no binding activity (Grenier *et al*., 1997; Gressent *et al*., 2003). Gressent *et al*., (2007) showed multiple insects regardless of sensitivity to have a high affinity PA1b binding site in their gut membranes. Based on the structural similarity of PA1b to ACTX, a spider toxin which blocks ion channels, it was hypothesised PA1b may have a similar function in the gut membrane (Jouvensal *et al*., 2003). The target site of PA1b was identified in a electrophysiological and pharmacological study as V-ATPase, specifically the V_0 subunit which is membrane bound and is available on the apical side of the gut membrane (Chouabe *et al*., 2011). The authors determined this through exposing patch clamped SF9 cells to PA1b and measuring the current across the cell membrane. Exposure to PA1b resulted membrane depolarisation in a dose-dependent manner. This provided evidence that PA1b was inhibiting a membrane channel or pump. Chouabe *et al*., (2011) found an almost identical response when they exposed the cells to bafilomycin, an irreversible specific

inhibitor of V-ATPase proton pumps, and further confirmed the target site of action by monitoring the ATPase activity of purified *Manduca sexta* V-ATPase in the presence of PA1b and found it to be reduced (Bowman *et al*., 1988; Chouabe *et al*., 2011). The specific binding site on the V-ATPase was elucidated using electron microscopy to show multiple binding sites on the *c-ring* of V_0 similarly to bafilomycin, mutations in the c-subunit which causes resistance to bafilomycin also conferred resistance to PA1b (Muench *et al*., 2014). The same study utilised radiolabelling and showed that the e-subunit also interacts with PA1b. It is hypothesised that PA1b prevents rotation of the *c*-ring by which is required for proton transport, by locking it to the static e-subunit (Figure 1.2). PA1b is the only known proteinaceous inhibitor of V-ATPase proton pumps. Inhibition of V-ATPases by PA1b in the sensitive *Sitophilus oryzae* strain WAA42, triggered apoptosis in effected cells (Eyraud *et al*., 2017). The function of PA1a is not known, it does not display insecticidal activity when exposed to SF9 cells, and it is suspected to aid in the correct folding of PA1b, as expression of PA1b in tobacco leaf tissue was poor (<2.4 μg g^{-1} leaf tissue) without the presence of a PA1a encoding sequence in the expression construct (Eyraud *et al*., 2013).

As mentioned previously, the primary amino acid structure and gene sequence of the PA1 gene cassette were elucidated in the mid-1980s (Gatehouse *et al*., 1985; Higgins *et al*., 1986). Eventually the tertiary structure of PA1b was determined using NMR modelling of PA1b purified from pea seed (Figure 1.3) (Jouvensal *et al*., 2003). The authors determined that this small peptide is a member of the inhibitor cysteine knot (ICK) family and displays all the structural characteristics of a knotting fold, specifically a triple-strand antiparallel β-sheet with a long flexible loop connecting the first and second strand, and the cysteine knot motif (Figure 1.3). This structure is undoubtedly related to its function, however the knottin protein group whilst similar in structure, are diverse in function and contain many biologically active proteins from plants, spiders, fungi and cone snails (Daly and Craik, 2011). These proteins are known to be responsible for a wide range of biological activity, but a large proportion of them inhibit voltage gated ion channels (Daly and Craik, 2011; Fletcher *et al*., 1997; Norton and Pallaghy, 1998). There is evidence to suggest that the ICK motif confers high proteolytic, thermal and chemical stability (Daly and Craik, 2011; Herzig and King, 2015). For example the ICK peptide Hv1a was shown to be stable from pH 1-8 and to resist degradation by proteinase K (Herzig and King, 2015). PA1b itself has been shown to be remarkably resistant to proteolytic degradation both to commercial enzymes and gut extracts from resistant insects (Gressent *et al*., 2007). Jouvensal *et al*., (2003) further suggested that PA1b had structural similarities to the toxin from

Hadronyche versuta ACTX-Hi, despite there being little sequence similarity. The authors further identified that PA1b as an amphipathic protein with an external hydrophobic face formed on the residues of L2, between the first and second β-sheets, this amphipathic nature is thought to be necessary for activity (Eyraud *et al*., 2013). A structural study went on to determine, through chemical mutagenesis, that the receptor binding and insecticidal activity were both dependent on a cluster of residues in L2, with binding activity being dependent on Phe-10, Arge-21, Ile-23 and Leu-27 (Da Silva *et al*., 2010).

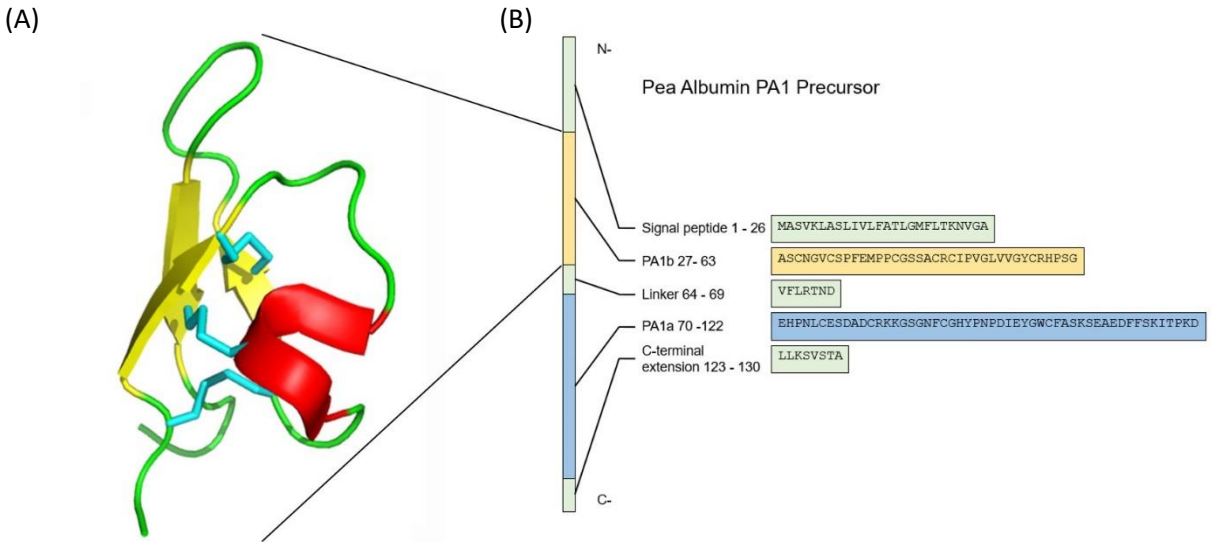


Figure 1.3 PA1b 3D - structure and the Pea albumin PA1 precursor. (A) Three dimensional structure of PA1b (Protein data Bank ID cod 1P8B) showing α- helix (red), β-sheet (yellow) and disulphide bridges (blue) (Jouvensal *et al.*, 2003). (B) The primary structure for the PA1 preproprotein (Gatehouse – Unpublished based on accession number: CAB82859.1).

1.6 *Galanthus nivalis* agglutinin (GNA)

Lectins are sugar binding molecules found in all plants, and some other organisms such as bacteria and were initially discovered in 1888 by Stillmark who discovered that an extract of castor beans hemagglutinated erythrocytes. This large group of heterogenous proteins are often categorised by their binding properties (Van Damme *et al*., 1998). Their function is primarily considered to be nitrogen-storage, and they bind specifically and reversibly to mono- or oligosaccharides (Peumans and Van Damme, 1995). Beyond nitrogen storage lectins are proposed to be involved in an array of biological functions ranging from conferring specific interactions between legumes and rhizobia to plant defence against insects or microorganisms (Gatehouse *et al*., 1995; Laus *et al*., 2006; Peumans and Van Damme, 1995).

The mannose specific Snowdrop lectin or *Galanthus nivalis* agglutinin (GNA) has been well characterised in the literature and was the first monocot mannose-binding lectin to be isolated (Van Damme *et al*., 1987). GNA exists *in vivo* as a 50 kDa homotetramer (Quaternary structure shown in Figure 1.4. Each of the four identical subunits has a mass of approximately 12 kDa and consists of 109 amino acid residues which form a β-barrel structure. Each subunit has three binding sites for mannose, resulting in a total of 12 binding sites per homotetramer (Hester *et al*., 1995). The mannose binding site of GNA monomers consists of four conserved amino acids (Gln89, Asp91, Asn93 and Tyr97). The specificity for α-1,3 or α-1,6 linked D-mannose is due to hydrogen bonding to Asp91 and Asn93 which is not possible with other sugars such as D-glucose (Barre *et al*., 2001). Functional recombinant GNA has been expressed in a range of heterologous expression systems including *Pichia pastoris, Nicotiana tabacum* and *Oryza sativa* (Raemaekers *et al*., 1999; Rao *et al*., 1998; Stoger *et al*., 1999).

Snowdrop lectin has been found to be somewhat toxic towards many important agricultural insect pests, across multiple orders, including the Hemiptera, Coleoptera and Lepidoptera. Toxicity of GNA has been shown at varying levels towards *Myzus persicae*, *Acrythosiphon pisum, Nilaparvata lugens,* (Powell *et al*., 1993; Rahbé *et al*., 1995a; Sauvion *et al*., 1996). Fitches, Gatehouse, and Gatehouse (1997) showed that larvae of the tomato moth *Lacanobia oleracea* fed on artificial diet containing GNA had significantly reduced biomass and extended development time as compared to controls. Expression of GNA in transgenic crops has also been shown to confer resistance, with varying degrees of success from decreased survival to decreased fecundity towards many different pests including multiple aphid species, plant hopper, *N. lugens* and lepidoptera *L*. *oleracea* (Gatehouse *et al*., 1996, 1997; Hilder *et al*., 1995; Rao *et al*., 1998; Stoger *et al*., 1999). Whilst GNA has some toxic activity towards insects importantly it does not display any toxic effects toward higher mammals, unlike other

insecticidal lectins such as the jackbean lectin Concanavalin A (ConA) (Oliveira *et al*., 1989; Poulsen *et al*., 2007).

The mechanism for GNA insecticidal activity is still not yet elucidated. GNA is highly resistant to gut proteolysis and has been shown to bind to proteins present in gut epithelial membrane, such as aminopeptidase, and accumulate in the gut in many insects (Cristofoletti *et al*., 2006; Fitches and Gatehouse, 1998; Powell *et al*., 1998; Yang *et al*., 2014). Further after feeding on GNA, the lectin can be detected in the haemolymph, fat bodies and ovarioles in *N. lugens* (Hemiptera). Similarly transport across the gut epithelium to the haemolymph following ingestion of GNA has also been demonstrated in Lepidopteran and Coleopteran pests (Fitches *et al*., 2001; Hogervorst *et al*., 2006). The mechanism of transport across the gut epithelium is known to be dependent upon binding to mannose residues and thought to occur via receptor meditated endocytosis (Bonning and Chougule, 2014). The ability of GNA, following ingestion, to be transported from the gut lumen to the haemolymph in a range of insects from different orders forms the basis for the development of novel insecticidal fusion proteins.

1.7 GNA fusion proteins as a novel insect control method

The ability of GNA to transport across the insect midgut barrier to the haemolymph after oral ingestion, opened up the possibility of utilising GNA to transport toxic molecules to their target site. The initial GNA fusion protein was a fusion of *Manducta sexta* allatostaitn (Manse-AS) and GNA, the fusion was found to inhibit feeding and development of *L. oleracea* when administered in diet, but it did not cause mortality (Fitches *et al*., 2001). The authors also detected high Manse-AS immunoreactivity in the haemolymph after feeding providing the first evidence that GNA can transport peptides across the gut epithelium. Building on previous success Fitches *et al*., (2004) fused a toxin from the spider *Segestria florentina* to GNA (SFI1/GNA), feeding of the fusion protein to *L. oleracea* resulted in 100 % mortality after six days whereas neither the toxin nor GNA alone displayed acute toxicity. Furthermore, intact SFI1/GNA was detected in the haemolymph after feeding. Various other toxins have been fused to GNA and tested against multiple insect species, and have all effectively increased the oral activity of neuropeptides against the plant hopper *N. lugens,* beetles *(T. castaneum* and *A. tumida*)*,* aphids (*A. pisum, M. persicae, S. avenae*) and lepidopteran larvae (*M. brassica*, *L. oleracea*, *S. littoralis*) (Down *et al*., 2006; Nakasu *et al*., 2014a; Powell *et al*., 2019; Trung *et al*., 2006; Wakefield *et al*., 2010; Yang, 2015; Yang *et al*., 2014). Expression of GNA fusions with the calcium ion channel blocker ω-ACTC-Hv1a, in transgenic Arabidopsis has also been shown to confer resistance to *M. persicae* as compared to wild type (Nakasu *et al*., 2014a).

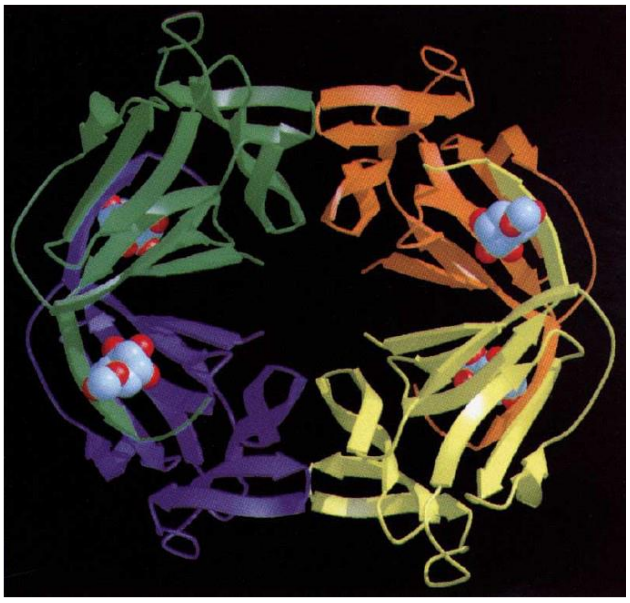


Figure 1.4 GNA homotetramer, quaternary structure. Individual subunits are displayed in different colours. The high affinity mannose binding sites are displayed on each subunit containing a bound sugar molecule. Source: Hester *et al*., (1995).

The same fusion proteins also showed no toxicity to honeybee (*A. mellifera*) larvae or adults and no detrimental effects on memory in adult bees (Nakasu *et al*., 2014a; Powell *et al*., 2019).

As PA1b has a target site in the insect gut, the anchoring properties of GNA are possibly more important than transport across the gut in regard to this project. Yang *et al*., (2014) demonstrated the ability of GNA to facilitate retention of toxin in the aphid gut through binding to the epithelial membrane. It is hypothesised that fusion of pea albumin to GNA could result in increased toxicity as a result of localisation to the gut membrane by binding to mannosylated proteins that line the gut epithelium.

1.8 *Pichia pastoris* as a protein expression system

1.8.1 History of *Pichia pastoris*

The methanotrophic yeast *P. pastoris* is a currently used as a major expression host in the production of heterologous proteins, throughout academia, and the biotechnology and pharmaceutical industries (Ahmad *et al*., 2014; Weinacker *et al*., 2014). The pervasiveness of this microorganism is due to several advantages it confers over other protein expression systems. These advantages include but are not limited to: the ability to form disulphide bridges within the cytoplasm, the ability to glycosylate recombinant proteins amongst other post-translational modifications not possible in prokaryotic cells such as *E. coli;* the expression levels are often significantly higher or at least comparable to other expression systems, and is scalable to industrial fermentation (Baumgartner *et al*., 2003). *Pichia pastoris* was initially isolated from the bark of a French chestnut tree and named *Zymosacchromyces pastoris* in 1920 by Alexandre Guilliermond. By 1995 another strain had been isolated, and these were designated to their own genus *Komagataella* (Phaff *et al*., 1956; Yamada *et al*., 1995). These two strains were then separated into two distinct species, *Komagatealla pastoris* and *K. phaffii* (Kurtzman, 2005). Both *K. pastoris and K. phaffii* strains are utilised for heterologous protein expression and are referred to as *P. pastoris* throughout the literature (Zahrl *et al*., 2017). Throughout this thesis the name *Pichia pastoris* is used, in keeping with published literature.

1.8.2 Promoter systems

There are multiple promoter systems that are available for use in *P. pastoris* protein expression. One of the most common and widely utilised is the AOX1 promoter, which utilises methanol to induce protein expression. Using an inducible promoter adds a degree of complication to the process of heterologous protein expression, as it requires a swap of carbon source from glycerol or glucose to methanol to induce protein expression (Tschopp *et al*., 1987).

The AOX1 promoter is often repressed by other commonly used carbon sources in benchtop fermentation (Inan and Meagher, 2001). Furthermore, methanol is a toxic and flammable compound, introducing safety concerns when scaling to large volume fermentations. In order to reduce complicating factors throughout this thesis the constitutive glyceraldehyde-3-phosphate dehydrogenase (GAP) promoter was used, which should continually produce protein throughout yeast lifecycle (Waterham *et al*., 1997).

1.8.3 Recombinant protein localisation

The localisation of protein after expression is an important factor in the production of recombinant protein. Secretion to the medium will significantly reduce the resources needed for purification by removing the need to lyse the cells to access the protein. *Pichia pastoris* has relatively few endogenous proteins which are secreted into the medium, with only 20 unique proteins identified in the secretome using LC-ESI-MS/MS and 21 using 2D gel electrophoresis, as compared to the *S. cerevisiae* secretome which contains 239 unique secreted proteins (Mattanovich *et al*., 2009; Smeekens *et al*., 2017). The combination of relatively simple methodology and relatively few detectable excreted proteins provide a good starting point for protein purification from culture supernatants.

The pGAP*α* vectors from Invitrogen, contain the S. cerevisiae α-factor signal sequence upstream of the protein coding sequence. The α-factor directs the heterologous protein to the culture medium after expression, and should be removed by post-translational processing before the protein is exported to the medium, and this has been demonstrated for the expression of GNA using transformed *P. pastoris cells*, (Brake *et al*., 1984; Raemaekers *et al*., 1999). The α-factor is processed in three steps. Initially the pre signal is cleaved in the endoplasmic reticulum, then Kex2 endopeptidase cleaves the pro sequence, between arginine and lysine, and finally, the glutamic acid alanine repeats should be removed by a Ste13 protein (Brake *et al*., 1984). On occasion, the α-factor is incorrectly processed, and the recombinant protein can have a glutamate-alanine repeat on the N-terminus, this does not generally affect activity.

1.9 Aims and objectives

This project explores the utilisation of GNA fusion protein technology to enhance the toxicity of pea albumin toward hemipteran pests. Various expression constructs, fermentation and purification strategies are investigated to determine the optimal method of production. Purified recombinant pea albumin – GNA fusion proteins were then tested for insecticidal activity towards three model aphid pests and compared to a recombinant version of pea albumin produced in *P. pastoris*. Finally, the reasons for the observed enhanced toxicity of the fusion protein as compared to pea albumin alone and the differential toxicity between *A. pisum* and *M. persicae* were investigated.

The objectives are as follows:

1. Evaluate the effect of excess glycerol in the fermentation medium on the recovery of recombinant GNA when purified using nickel affinity chromatography
2. To evaluate samples of PA1b/GNA produced previously to confirm initial findings. Then subsequently produce a clean sample of a PA1b/GNA fusion protein to determine if the acute toxicity initially observed was due to poor sample quality or due to high activity of the fusion protein.
3. Design new fusion protein constructs with the aim of improving expression levels and recovery of purified recombinant fusion proteins using a *P. pastoris* expression system.
4. Evaluate the insecticidal effects of pea albumin fusion proteins and determine why fusion to GNA may enhances activity compared to recombinant pea albumin.

Chapter 2 |Materials and Methods

2.1 Recipes of commonly used media, buffers and reagents

2.1.1 Bacterial culture medium

LSLB Broth:	1% (w/v) NaCl, 1% (w/v) peptone, 0.5% (w/v) yeast extract, dissolved in distilled water.
LSLB Agar:	1.5% agar added to LB Broth

2.1.2 Yeast culture medium

YPG broth:	1% (w/v) yeast agar, 2% (w/v) tryptone, 4% (w/v) yeast extract, dissolved in distilled water.
YPG Agar:	1.5% Bacto agar added to YPG medium.
Fermentation medium:	Basal salt medium as per Higgins and Cregg (1998). Yeast peptone glycerol medium (YPG) Buffered glycerol complex medium (BMGY)

2.1.3 Agarose gel electrophoresis:

TAE (50X):	2M Tris/Acetic acid pH 7.7, 50mM EDTA.
DNA Loading buffer:	30 % (v/v) glycerol, 0.25 % (w/v) bromophenol blue, sterile distilled water.
Agarose gel:	1X TAE, 0.8 % (w/v) agarose

2.1.4 Protein gel electrophoresis (SDS-PAGE):

5X SDS sample buffer:	0.5M Tris/HCl (pH 6.8), 50% (v/v) glycerol, 5% (w/v) SDS, 0.005% (w/v) bromophenol blue.
Acrylamide:	40% (w/v) acrylamide: 0.8% (w/v) bis-acrylamide stock solution (37.5:1) (Sigma-Aldrich, St. Louis, USA).
Resolving buffer:	3.0 M Tris/HCl pH 8.8.
Stacking buffer:	0.5 M Tris/HCl pH 6.8.
Reservoir buffer (10X):	0.25M Tris/HCl pH 8.3, 1.92 M Glycine, 1% (w/v) SDS.
(CBB) Stain:	40% (v/v) methanol, 7% (v/v) glacial acetic acid, 0.05% Coomassie Brilliant Blue (CBB).
Destain:	40% (v/v) methanol, 7% (v/v) glacial acetic acid.
Stacking gel mixture:	2.5% (v/v) Acrylamide (37.5:1 acrylamide:bis-acrylamide; Sigma-Aldrich, St. Louis, USA), 125mM Tris-HCl (pH 6.8), 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate, 0.0075% (v/v) *N,N,N',N'*-tetramethhylethylenediamine (TEMED).

Resolving gel mixture: 12.5% or 15% or 17.5% (v/v) Acrylamide, 375mM Tris-HCl (pH 8.8), 0.1% (w/v) SDS, 0.075 (w/v) ammonium persulphate, 0.05% (v/v) *N,N,N',N'*-tetramethylethylenediamine (TEMED).

2.1.5 Protein molecular weight marker:

Pierce™ Unstained Portein Marker:	Mass
	116.0 kDa
	66.2 kDa
	45.0 kDa
	35.0 kDa
	25.0 kDa
	18.4 kDa
	14.4 kDa

2.1.6 Western blotting:

Transfer buffer: 48 mM Tris/HCl, 39 mM glycine, 20% (v/v) methanol, 0.0375% SDS (pH 9.2).

Ponceau stain: 0.1% Ponceau S, 5% (v/v) acetic acid, prepared in distilled water.

PBS (10X): 0.015 M KH_2PO_4, 0.08 M Na_2HPO_4, 1.37 M NaCl in distilled water.

Blocking solution: 5% (w/v) Non-fat milk powder, 1X PBS, 0.1% (v/v) Tween-20.

Anti-Sera solution: 5% (w/v) Non-fat milk powder, 1X PBS, 0.1% (v/v) Tween-20.

PBST: 1X PBS, 0.1% (v/v) Tween-20.

2.1.7 Chemiluminesent detection reagents:

Solution A: 100 mM Tris/HCl pH 8.0, 0.2 mM coumaric acid, 1.25 mM luminol in 50 mL distilled water.

Solution B: 10 % (v/v) H_2O_2 (30% solution) in distilled water.

2.1.8 Schiff – periodic acid staining

Solution A: 1.0 % (v/v) periodic acid, 3 % (v/v) acetic acid

Solution B: 0. 1 % (w/v) Sodium metabisulphite in 10 mM HCl

2.1.9 DNA molecular weight marker:

Bioline Hyperladder 1kb	Size:	ng/Band:
	10037	100
	8000	80
	6000	60
	5000	50
	4000	40
	3000	30
	2500	25
	2000	20
	150/1517	15/15
	1000	100
	800	80
	600	60
	400	40
	200	20

2.1.10 Nickel affinity chromatography buffers

Binding buffer (4x):	80 mM sodium phosphate, 2 M NaCl, pH 7.4
Binding buffer (1x):	20 mM sodium phosphate, 0.5 M NaCl, pH 7.4
Wash buffer:	20 mM sodium phosphate, 0.5 M NaCl, 10 mM imidazole, pH 7.4
Elution buffer:	20 mM sodium phosphate, 0.5 M NaCl, 200 mM imidazole, pH 7.4

2.1.11 Strep-tactin chromatography buffers

Binding buffer (4x):	80 mM sodium phosphate, 1.1 M NaCl, 24 mM potassium chloride, pH 7.4
Binding buffer (1x):	20 mM sodium phosphate, 280 mM NaCl, 6mM potassium phosphate, pH 7.4
Elution buffer:	20 mM sodium phosphate, 280 mM NaCl, 6mM potassium phosphate, 2.5 mM desthiobiotin, pH 7.4
Regeneration buffer:	0.5 M NaOH

2.1.12 Yeast lysis buffer

Yeast lysis buffer:	2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0)

2.2 Standard molecular biological techniques

2.2.1 Oligonucleotide primers

Oligonucleotide primers were produced by Integrated DNA Technologies (IDT, Coralville, USA). The primers were resuspended in sterile distilled water to a concentration of 100 µM. Before use primers were diluted using sterile distilled water to final concentrations of 20 µM.

2.2.2 Amplification of DNA by polymerase chain reaction (PCR)

PCR was performed using a Thermo Scientific Hybaid PCR Sprint thermal cycler. PCR reactions, between 25 and 50 µl, were set up on ice, in 0.2 mL PCR tubes. A reaction of 50µl generally consisted, DNA template (5-100ng), 10 µL Hifi polymerase buffer (PCR biosystems, London, UK), 1 µM of each the respective forward and reverse primer and 0.5 µL of Hifi polymerase (PCR biosystems, London, UK), then made up to 50 µL with sterile distilled water. For multiple reactions a master mix containing all reagents common to all reactions was prepared and dispensed into individual tubes, where upon the differing reagents were added. Typically, PCR followed stages of initial denaturation at 95 °C for 5 minutes, then 25 – 35 cycles of denaturation at 95 °C for 30 seconds, annealing at 55 °C for 30 seconds and extension at 72 °C for 30 seconds to one minute depending on template length. Finally, a final extension stage at 72 °C lasted for 7 minutes.

2.2.3 Agarose gel electrophoresis

DNA was separated using agarose (Sigma-Aldrich, St. Louis, USA) gels of 1-2% (w/v) in a 1X TAE buffer containing 0.5 µg mL^{-1} Ethidium Bromide to facilitate visualization with UV light (UVB, 300 nm). Samples were prepared before loading to contain 1X DNA loading buffer. The gels were then run at 50-100 V in 1X TAE buffer. The gels were then photographed using UV (300 nm) light.

2.2.4 DNA extraction from gel

After identification of DNA fragments under UV (300 nm) light, DNA bands of the appropriate mass were excised from the agarose gel using a clean scalpel, whilst minimising exposure to UV (300 nm) light. The gel sliced was weighed on a fine balance. DNA fragments were extracted using the QIAquick Gel Extraction kit, following the manufacturers protocols and eluted in 50 µL of nuclease-free water.

2.2.5 Restriction endonuclease digestion of DNA

Restriction endonuclease digests were carried out using commercially available restriction endonuclease enzymes with manufacturer recommended buffers. Typically, 2 μg of plasmid DNA, 4 μL of 10 X reaction buffer, 2 μL of the appropriate enzyme and the reaction was made up to 40 μL with NFW. The digestions were incubated at 37°C overnight for complete digestion.

2.2.6 DNA ligation

DNA ligation reactions were carried out using a commercially available T4 DNA ligase (Promega, Madison, USA), following the manufacturer recommended protocols. Reactions consisted of 1 μL of 10 X T4 buffer, 1 μL of T4 ligase, appropriately restricted insert DNA and vector DNA in a 3:1 ratio, the reaction was made up to 10 μL with sterile distilled water. Reactions were incubated on the bench for 1 hour followed by overnight at 4 °C.

2.2.7 Bacterial culture

Liquid bacterial cultures were grown in 10 mL Low Salt Luria-Bertani (LSLB) medium using single colonies picked from LB-agar plates. Liquid cultures were grown at 37 °C overnight on a rotary shaker at 220 rpm. If appropriate antibiotics were added to culture medium to facilitate selection. When cultured on solid LSLB-agar plates, antibiotics were added when the medium had cooled to approximately 50°C. LSLB medium were used with Zeocin (Invitrogen, Carlsbad, USA) for antibiotic selection.

2.2.8 Electro-competent cells

Electro-competent *E. coli* TOP10 cells were prepared as described in (Fitches *et al*., 2004), and were used for transformation.

2.2.9 Transformation of *E. coli*

Electro-competent cells were transformed as per a standard protocol outlined in Sambrook and Russel (2001). 50μl of electro-competent cells were mixed with between 0.5 and 1.0μl of DNA ligation reaction mixture, and transformed under the following conditions: Biorad Electroporator, electrical pulse set to 25μF, capacitance set to 2.5kV and 200ohm resistance. The above conditions were applied to the cells for 4-5 milliseconds with field strength of 12.5kV/cm. The transformation mix was incubated for 1 hour at 37°C and then plated onto the

respective medium containing antibiotic. The plates were then incubated overnight at 37°C. Colonies were screed using colony PCR and positive results were sent for sanger sequencing.

2.2.10 Colony PCR

Colony PCR reactions were 20 µL and consisted of 2 µL of 10 x PCRbio Classic Buffer (PCR biosystems, London, UK) with 30 mM $MgCl_2$, 0.5 µL of 25 µM dNTP solution, 0.5 µL of each 20 µM specific forward and reverse primers, 1 µL of Classic Taq polymerase and then made up to 20 µL with sterile distilled water. If numerous reactions were required a master mix was made up to a volume of n + 1 and aliquoted into 0.2-mL tubes. A bacterial colony was picked from the transformation plate using a sterile pipette tip and transferred into individual PCR tubes, and vortexed. Colony PCR utilised a cycle of 95 °C for 10 minutes, then 25 cycles of denaturation at 95 °C for 30 seconds, annealing at 55 °C for 30 seconds and extension at 72 °C for 30 seconds to one minute depending on template length. Finally, a final extension stage at 72 °C lasted for 7 minutes.

2.2.11 Isolation of plasmid DNA

Clones were incubated in 10 mL of LSLB with appropriate antibiotic selection at 37 °C overnight, centrifuged, and plasmid DNA was isolated from the cell pellet using the Wizard *Plus* SV miniprep DNA purification kit (Promega, Madison, USA) as per the manufacturer's instructions. Plasmid DNA was eluted in 100 µL of nuclease free water and stored at -20°C.

2.2.12 DNA sequencing

DNA sequencing was carried out by the DNA sequencing service (DBS Genomics) at the Department of Biosciences at Durham University. Sanger sequencing was performed using Applied Biosystems ABI Prisim 3730 automated DNA sequencers. Plasmid specific primers were used in all DNA sequencing reactions. SerialCloner (Version 2.6.1) was used to view and analyse the sequence data.

2.2.13 Linearization of plasmid DNA

To prepare DNA for transformation into *P. pastoris* plasmid DNA was linearized using AvrII. Plasmid isolation of 40 mL of LSLB was pooled and restricted with AvrII. The reaction consisted of 200 µL plasmid, 10 µL AvrII, 25 µL 10 C CutSmart Buffer and 15 µL of sterile distilled water. Linearization reactions were incubated over night at 37 °C.

2.2.14 Ethanol precipitation of DNA

The volume of DNA solution was checked using a pipette and 1/10th volume of 3 M sodium acetate was added along with a double volume of absolute ethanol. This was then incubated overnight at -20°C. The solution was then centrifuged at 14,000 rpm for 15 minutes at 4°C. The pellet was washed and re-pelleted with 70% ethanol twice, then air dried and re-suspended in an appropriate volume of sterile distilled water.

2.2.15 Sodium dodecyl sulphate-polyacrylamide gel glectrophoresis (SDS-PAGE)

For the separation of protein, the samples and a molecular weight marker (Thermo Pierce unstained protein MW marker) were denatured and reduced before loading by adding concentrated sample loading buffer (5X) enough sample to reach 1X and boiling for 10 minutes at 100°C. The proteins were then separated according to weight by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) by the protocols outlined in Laemmli *et al*., (1970). Gels (9x10cm) were run in 1X reservoir buffer at constant voltage (100-150V) in ATTO AE450 gel tanks. Following electrophoresis, the gels were either stained or transferred to nitrocellulose membrane for western blotting.

2.2.16 Coomassie blue staining and de-staining of SDS-PAGE gels

Proteins were visualised by staining with CBB stain for a minimum of 3 hours, often overnight, followed by immersion in destain solution until the background was clear. Both processes were performed at room temperature with gentle shaking.

2.2.17 Western blotting

Proteins were blotted on to nitrocellulose (Hybond ECL, Amersham, Amersham, UK) membrane following SDS-PAGE by semi-dry electro-blotting. The nitrocellulose membrane and 3MM blotting paper (Whatman) were cut to the same dimension as the gels and soaked in semi-dry transfer buffer for up to 1 hour before blotting. Three layers of blotting paper were placed on the bottom electrode of the blotter followed by the nitrocellulose membrane. The acrylamide gel was placed on top of the membrane followed by three more sheets of blotting paper, air bubbles were removed with a plastic roller. The electro-blotting was performed on an ATTO AE-6675 blotting apparatus. The conditions of the transfer were 150mA (2.0 mA/cm^2) constant current for 60 minutes. Transfer was confirmed by staining with Ponceau S stain. The protein marker bands were marked on the membrane with pencil, and the membrane was then destained with distilled water.

2.2.18 Chemiluminescent detection of proteins

In order to block the non-specific protein binding sites, membranes were incubated in 40 mL blocking solution and agitated gently for 20 minutes, three times. Following blocking the membranes were incubated overnight at 4°C with 5 mL blocking solution containing primary antibody (Anti-His or Anti-GNA), with a ratio of 1:2500 unless otherwise specified. Unbound primary antibody was removed by three ten-minute washes with antisera at room temperature. The membrane was then transferred to antisera containing appropriate secondary antibody (Goat Anti Mouse IgG(H+L)-HRP conjugate) in the same ratio as the primary antibody and incubated for 1-2 hours at room temperature. Membranes were then washed three times with 1 X PBST, 10 minutes per wash. 3 x 10 minute washes with distilled water were used to remove excess Tween 20. Solution A and B were prepared as described in the recipe. 5 mL of Solution A was mixed with 15 μl of solution B immediately before detection. Specific antibody binding was visualised by exposing membranes to photosensitive film (SLS). Exposed films were washed and develop on an X-ograph Imaging Systems Compact X4 automatic developer.

2.3 Transformation of *Pichia pastoris*

Plasmid DNA was collected from clones verified for the correct sequence through a mini-prep kit. Plasmid DNA from 3x10 mL cultures were pooled and linearized with the restriction enzyme BlnI or AvrII overnight at 37°C, following the standard restriction digestion protocol. Linearization was confirmed on a 1 % agarose gel with undigested plasmid as a control. The digested DNA was precipitated using ethanol and suspended in 20μl sterile distilled water and transformed into chemi-competent cells of proteases deficient *P. pastoris* (SMD1168; Invitrogen, Carlsbad USA) and X33 using the *Pichia* EasyComp transformation kit as per manufacturer's instructions. Following transformation cells were incubated for two hours at 30 °C in yeast peptone glycerol medium (YPG) to allow expression of Zeocin resistance. The transformants were then selected on (25 μg mL^{-1}) Zeocin in YPG-agar plates by growing for 2-3 days.

2.4 Screening of *Pichia pastoris* transformants

Transformants were grown in small baffled flasks containing liquid YPG (10 mL) and Zeocin (100 μg mL^{-1}) for 48-72 hours at 30°C with shaking (200 rpm). The cultures were centrifuged at 6000 rpm at room temperature and 25 μl was used for SDS-PAGE followed by western blotting with appropriate antibodies.

2.5 Yeast genomic DNA extraction

Yeast genomic DNA extraction was performed based on Harju, Fedosyuk, and Peterson (2004). Single colonies were picked from the transformation plates and inoculated into 10 ml of YPG with 100 μg mL^{-1} zeocin and incubated at 30 °C for 48 hours. One mL of each culture was pelleted in a microcentrifuge tube and the cell pellets were resuspended in 200 μL of lysis buffer. The tubes were placed in liquid nitrogen for two minutes then placed in a 100 °C heat block for 1 minute to thaw quickly. The process was repeated five times, and the tubes were vortexed vigorously for 30 seconds. 200 μL of chloroform was added and the tubes were vortexed for 2 minutes and then centrifuged 3 minutes at room temperature at 20,000 × g. The aqueous layer was transferred to a tube containing 400 μL of ice-cold 100% ethanol. The samples were allowed to precipitate at room temperature and then centrifuged 5 minutes at room temperature at 20,000 × g. Supernatants were discarded, and DNA pellets were washed with 500 μL of 70% ethanol followed by air drying at room temperature. DNA was resuspended in 20 μL nuclease free water.

2.6 Expression of fusion proteins through benchtop fermentation

A clone which showed good levels of expression blot was selected and grown on fresh YPD-agar plates containing Zeocin (25 μg mL^{-1}), along with 3x 100 mL YPG liquid cultures (in 250 mL baffled flasks) were inoculated from a master plate to initiate the starter culture. The culture was grown for 2-3 days at 30 °C whilst shaking at 200 rpm. A 7 L benchtop laboratory fermenter (Applikon) was used to grow the selected *P. pastoris* clone, to produce the recombinant protein, the methodology used was similar to that used in Fitches *et al*., (2004). 2.5 litres of basal salt medium (BSM) were prepared, as in Higgins and Cregg (1998). Alternate media used were half concentrated YPG medium or Buffered Glycerol-complex medium (BMGY) as in Potgieter *et al*., (2009). The medium and fermenter were sterilised by autoclaving. Sterile BSM was supplemented with PTM1 salts (Cino, 1999). The fermentation was carried out at 30°C with constant agitation at pH 4.5-5 (YPG was maintained at pH 6.5 and BMGY was maintained at pH 6) and constant dissolved oxygen level of 30 %. When cell density reached a point where 30% dissolved oxygen could not be maintained by agitation, a constant feed of glycerol at 50% (v/v) was initiated (5-10 mL/hour). When the glycerol feed was depleted the run was continued for 2 – 3 hours to allow utilisation of the glycerol then the run was stopped.

2.7 Processing of fermenter supernatant

The culture from the fermenter was centrifuged at 800rpm at 4°C for 30 minutes. The supernatant was filtered through 2.7 μM, followed by 0.7 μM glass microfiber filters using a vacuum manifold and then diluted with desired binding buffer for loading onto nickel affinity columns (Ni-NTA), specifically HisTrap FF columns (GE Healthcare), or StrepTrap HP columns (GE Healthcare). Purification of protein was performed using nickel affinity column chromatography or by Strep-Tactin affinity chromatography. An aliquot of supernatant was saved prior to dilution as load control for analysis by SDS-PAGE or western blotting.

2.8 Nickel affinity chromatography

2.8.1 Standard protocol

Ni-NTA columns (HisTrap FF, GE Healthcare) were washed with dH_2O to remove ethanol from storage, then equilibrated with 10 column volumes (CV) of 1 x binding buffer using peristaltic pumps. Culture supernatant was then loaded onto the column continuously using a peristaltic pump set to 5 mL min^{-1} for a minimum of 2 hours up to overnight. After loading a sample was saved of the 'flow through' for analysis by SDS-PAGE or western blotting. The ouput tubing was attached to a UV (280 nm) absorbance monitor connected to a chart recorder so absorbance (used as a proxy for protein content) could be monitored. The column was then washed with 10 CV of 1X binding buffer (W1) or the chart recorder reached baseline, which ever occurred second, the first 5 mL of W1 was collected for analysis on gel. The columns were subsequently washed with 10 CV wash buffer (W2) to remove non-specifically bound or weakly bound proteins, the peak fraction of W2 was collected for analysis on gel. Finally, the protein was eluted using the elution buffer and the peak fraction (on chart recorder) was collected. The column was then washed with binding buffer and loaded again.

2.8.2 Purification of PA1b/GNA

Initially the culture supernatant was dialysed against distilled water to remove salts this was then lyophilised as described in section 2.10. The lyophilised powder was then resuspended in equivalent volumes of nickel affinity binding buffer with 6 M urea. This was then purified as above. Then dialysed against a decreasing gradient of urea concentration then lyophilised.

2.9 Strep-tactin affinity chromatography

Strep-tactin affinity columns (StrepTrap HP, GE Healthchare) were washed with 10 CV of binding buffer (1x) to remove ethanol from storage and equilibrate the column. Culture

supernatant was then loaded onto the column continuously using a peristaltic pump set to 5 mL min^{-1} for a minimum of 2 hours up to overnight. After loading a sample of supernatant was saved and designated as the flow through. connected to a chart recorder so absorbance (used as a proxy for protein content) could be monitored. The column was then washed with 10 CV of 1X binding buffer (W1) or the chart recorder reached baseline, which ever occurred second, the first 5 mL of W1 was collected for analysis on gel. Finally, the protein was eluted using the elution buffer and the peak fraction (on chart recorder) was collected. The columns were regenerated with 3 CV of distilled water followed by 3 CV of regeneration buffer and then 3 CV of distilled water. Columns were then re-equilibrated with 5 CV of binding buffer and then the supernatant was loaded again.

2.10 Dialysis and lyophilisation of protein samples

2.10.1 Preparation of dialysis tubing

The dialysis tubing was prepared by boiling in a solution of 2 % (w/v) sodium bicarbonate and 1mM EDTA (pH 8.0) for 20 minutes. The tubing was then rinsed thoroughly and stored in 20 % ethanol to prevent contamination.

2.10.2 Dialysis and lyophilisation

After elution the peak fractions were pooled and poured into dialysis tubing with either 12-14 kDa pores or 2-3 kDa pores, pore size was determined by the mass of the recombinant protein. The filled tubing was placed in a 15 L bucket of distilled water with a small quantity of 50 mM ammonium hydrogen bicarbonate and stirred slowly. The dialysis was carried out at 4°C overnight with constant stirring. The water was changed 4 times, once after the overnight step then every hour. to give the 100:1 ratio of water to sample. Post-dialysis the samples were frozen on the walls of freeze-drying flasks by shelling in liquid nitrogen and lyophilized on a vacuum freeze-dryer until only a dry powder remained, approximately 24-48 hours.

2.11 Quantification of recombinant proteins

Following lyophilisation, the protein powder was weighed and stored in a 50 mL plastic falcon tube. 1 mg of the lyophilised powder was resuspended in 200 µL of PBS, to give a 5 mg mL^{-1} (powder) solution. This solution was then run on 17.5 % polyacrylamide gels by SDS-PAGE and were visually compared to protein standards of a known concentration. This allows estimation of percentage protein per mg of lyophilised powder.

2.12 Agglutination activity assay

Agglutination assays were carried out with rabbit erythrocytes in a round bottom (U) clear microtiter plate. All lectin solutions for testing were made up to 0.1 mg mL^{-1} of lectin content along with a positive control, commercially purified GNA (Sigma-Aldrich, St. Louis, USA). 40 μL of rabbit blood stock was mixed with 1 mL of 1 X PBS and centrifuged to bring the cells to the bottom. The supernatant was discarded, and the pellet was resuspended in 1 mL 1 X PBS to make a 2 % rabbit erythrocyte solution. 50 μL of 1 X PBS was added to each well of the microtiter plate. 50 μL of lectin solution was pipetted into individual wells in successive rows then 2-fold serially diluted in successive rows. 50 μL of the 2 % erythrocyte solution is added to each well and mixed gently with the pipette. The plate was then incubated for 2 hours at RT. The threshold concentration required for agglutination was assessed visually. Negative agglutination results in a round pellet of cells and positive agglutination shows an even red colour.

2.13 Schiff-periodic acid nitrocellulose membrane staining

Proteins were transferred to a nitrocellulose membrane as described in section 2.2.17. The nitrocellulose membrane was then stained with Ponceau to visualise the molecular weight marker, which was marked on the membrane with pencil. The Ponceau was detained with water. The nitrocellulose membrane was then washed with distilled water for 5 minutes and subsequently incubated in solution A for 30 minutes. The membrane was further washed twice for 5 minutes in distilled water followed by two 5minute washes in solution B. From this point on the membrane must be kept in the dark. The membrane was incubated for 15 minutes with Schiff's reagent (Sigma-Aldrich, St. Louis, USA). The membrane was then washed twice more with solution B whilst remaining in the dark. The membrane was then air dried overnight and imaged on a scanner.

2.14 LC-MS analysis

For LC-MS analysis proteins in excised gel bands were digested with chymotrypsin or trypsin proteases. Chymotrypsin digests contained ProteaseMAX surfactant (Promega, Madison, USA) and were performed as described in the protocol for this product (using 30 ng/μl enzyme). Trypsin and chymotrypsin digestions were for 16 h 30 °C. Gel supernatants containing peptides were made to 0.5% trifluoroacetic acid, lyophilised and re-suspended in 16 μl 0.1% TFA prior to de-salting with a C18 ZipTip (Millipore). Samples were lyophilized and made to 12 μl with 0.1% TFA. LC-MS analysis of 5 μl aliquots was performed with a Sciex TripleTOF 6600 mass spectrometer coupled to an ekspert™ nanoLC 425 with low micro gradient flow module

(Eksigent) via a DuoSpray source (Sciex). Peptides were separated on a C18 reverse phase column using a gradient of acetonitrile in 0.1% formic acid. MS-MS acquisition used precursor-ion scans (400 to 1600 m/z) of 250 msec followed by CID-fragmentation of up to 10 multiply-charged ions, with an MS/MS spectrum (m/z 100-1600) acquisition time of 50 msec for each selected ion. MS-Analyst software version 1.7.1 (SCIEX) was used to acquire all MS and MS/MS data, MSConvert (ProteoWizard software suite) was used to generate peak lists, and protein identification used Peaks Studio 8.5 in conjunction with a database containing known *Pichia pastoris* protein sequences, known proteomic-experiment contaminants and the expressed proteins reported here.

2.15 Statistical analysis

All statistical analysis was performed using Graphpad Prism 8. Survival data was analysed using Log-rank (Kaplan-Meier) survival analysis. T-tests were performed using Welch's correction which does not assume equal variance in the samples. LC_{50} values were calculated by performing a non-linear regression based on the log of the concentration of the toxin versus the corrected moralities. P values of less than 0.05 are taken to be significant throughout this thesis.

2.16 Artificial rearing Pea aphid (*Acyrthosiphon pisum*), Peach-potato aphid (*Myzus persicae)* and Grain aphid (*Sitobion avenae)*

Parthenogenic females of the pea aphid *A. pisum* were maintained on bean seedlings, (*Vicia faba*). Whilst parthenogenic females of *M. persicae* were maintained on Chinese cabbage seedlings, (*Brassica rapa*). Parthenogenic females of *S. avenae* were maintained on oat seedlings (*Avena sativa*) Plants were maintained at 20°C in a long-day light regime of 16h of light followed by 8h of dark.

In artificial feeding trails, parthenogenetic adult females were transferred to sterile liquid diet containing vitamins, minerals 150mM amino acids and 500mM sucrose. After 24 hours on artificial liquid diet the day-1 neonate aphids (day 1 and day 2 for *M. persicae* and *S. avenae*) are used in feeding bioassays. Feeding chambers were maintained under the same environmental conditions as the cultures maintained on plants.

2.16.1 Aphid liquid diet (Kunkel 1976; Prosser and Douglas, 1992)

Stock solutions of amino acids, vitamins and minerals were prepared in advance and stored at -20°C until required.

Vitamin Stock:

Vitamin	Weight in mg
Biotin	0.1
Patnothenate	5
Folic acid	2
Nicotinic acid	10
Pyridoxine	2.5
Thiamine	2.5
Choline	50
Myo-inositol	50

Dissolved in 5 mL distilled water and separated into 500µl aliquots.

Amino acid stock (150mM):

Amino Acid:	Weight in mg
Alanine	50.8
Asparagine	213.9
Aspartate/Aspartic acid	189.7
Cysteine	42.5
Glutamic acid	123.6
Glutamine	241.1
Glycine	9.0
Proline	65.6
Serine	59.9
Tyrosine	10.9
Arginine	300.2
Histidine	182.4
Isoleucine	114.1
Leucine	114.1
Lysine	158.9
Methionine	43.5
Phenylalanine	47.1
Threonine	103.6

Tryptophan	58.2
Valine	101.9

Dissolved in 50 mL distilled water and separated into 5 mL aliquots.

Mineral stock:

Mineral	Weight in mg
$FeCl_3.6H_2O$	11
$CuCl_2.4H_2O$	2
$MnCl_2.6H_2O$	4
$ZnSO_4$	17

Dissolved in 10 mL distilled water and separated into 500µl aliquots.

When the diet was to be prepared, the stock solutions of amino acids, minerals and vitamins were thawed. 5 mL of amino acid stock was mixed with 100µl mineral stock and 500µl vitamin stock. The sucrose mix was prepared fresh on the day as follows and added to the above mixture:

Ascorbic acid	10mg
Citric acid	1mg
$MgSO_4.4H_2O$	20mg
Sucrose	1.7g (500mM)

150mg $K_2HPO_4.3H_2O$ was dissolved in 1 mL sterile distilled water and was added to the diet. The pH of the mixture was confirmed at approximately 7.0, using indicator paper. Volume was made up to 8 mL by adding distilled water, if needed. The diet was then filter sterilised using a 0.2 µm filter and stored at -20 °C until required.

When ready for use 20 µl of PBS or PBS containing recombinant protein was added to 80 µl, to a total volume of 100 µl of mixed diet.

2.17 Aphid oral toxicity assays

The protein of interest dissolved in PBS was incorporated into the aphid diet to a number of different final dietary concentrations. A layer of parafilm was stretched over a 3 cm diameter feeding chamber, 25 µl of diet with the protein of interest was placed on top and another layer of parafilm was stretched over the top to sandwich the diet, simulating a leaf. The nymphs of the species of interest were removed from the artificial diet where they were produced and gently placed onto the parafilm of an upturned feeding chamber using a clean, soft paintbrush. The aphids were given time to settle, then the chamber was righted making sure no gaps were

present around the base of the feeding chamber. Survival was assessed daily throughout the assay and the feeding chamber and diet replaced every 48 hours. The assay was ended when one condition began producing nymphs.

2.18 Aphid fecundity assays

Using the same aphis as the oral toxicity assay, any treatments which had surviving adult aphids were maintained, nymphs were counted daily and diet was changed every 48 hours as before until day 16. The number of adult aphids at the beginning were recorded and the number which developed wings was recorded. Average nymphs per non-winged adult was calculated and used as a measure of fecundity.

2.19 Aphid free choice feeding assays

Free choice feeding assays were performed similarly to the oral toxicity assays. PAF/GNA, PAF or GNA at equal molar concentrations (16 nmol mL^{-1}) and a control diet were placed on 3 cm parafilm covered feeding chambers, so the diets did not mix. Twenty nymphs aged day 1 *A. pisum* or day 1 and 2 *M. persicae* were placed in-between the two diets gently using a soft paint brush. After 24 and 48 hours the number of nymphs present on each diet was counted only nymphs within the visible diet circle were feeding. Aphids which were not feeding were discarded. Welch's t-test was used for statistical analysis of 6 reps for each protein and species.

2.20 Aphid total soluble protein extraction

Aphid total soluble protein extractions were performed by placing aphids into a microcentrifuge tube and snap freezing aphids by immersion of the tube in liquid nitrogen. After freezing 2.5 µL if ice-cold PBS was added to the tube per aphid, and they were homogenised by hand using a mini pestle, then incubated on 1ce for 15 minutes. The homogenates were then centrifuged at 4 °C for 20 minutes at 16000 x g, the supernatant was collected as the soluble protein fraction.

2.21 Aphid gut dissection

Aphid guts were dissected by submerging live adults in ice-cold PBS, fine tip forceps were used to grasp the head and abdomen of the insect. Gently the head was separated from the body, leaving only the gut tissue connecting the head and main body. The gut was gently separated from the head and the body with the fine forceps. The same process was performed with the body by grasping the hindgut and gently pulling away from the body.

2.22 Conjugation of recombinant proteins with fluorescein isothiocyanate.

Recombinant PAF/GNA, PAF, GNA and ovalbumin (control) were fluorescently labelled with a 2:1 molar excess of fluorescein isothiocyanate (FITC, Sigma-Aldrich, St. Louis, USA). Recombinant proteins were resuspended in 1 mL in 500 mM carbonate buffer pH 9.0 to a concentration of 2 mg mL^{-1} and centrifuged to bring any insoluble particulate to the bottom of the tube. The supernatant was transferred to a clean microcentrifuge tube then incubated with 2 times the proteins molar concentration of FITC (1 mg mL^{-1} DMSO) with rotation for 4 h at RT, under dark conditions. Samples were then dialysed in darkness against phosphate buffered saline (PBS pH 7.4) at 4 °C at a ratio of 200:1 (buffer:sample) to remove any excess unbound FITC.

2.23 Fluorescence microscopy

The gut was dissected from *A. pisum* as described in section 2.21, after day 5 nymphs had fed for 24 hours on artificial diet containing either 8 nmol mL^{-1} of either PAF/GNA, PAF, GNA, ovalbumin (5 nmol mL^{-1}) or control (no protein) diet with propidium iodide (100 ng μL^{-1}, counterstain) and subsequently chased with artificial diet and propidium iodide for either 24 or 48 hours. Following dissection guts were mounted directly onto glass slides in Vectashield H-100 and overlaid with glass coverslips. Guts were visualised using a Zeiss 880 confocal laser scanning microscope. Setting were as follows: Objective Plan-Apochormat 20x/0.8 M27; FITC excitation and image capture: 488 nm laser (0.5 % power), wavelengths collected between 493-572 nm, gain 575; Propidium iodide excitation and image capture: 543 nm laser (1.5 % power), wavelengths collected between 584-735 nm, gain 800.

2.24 Measuring average fluorescence using FIJI

To quantify the fluorescence in each treatment a minimum of 6 guts from each condition were imaged utilising the same laser and image settings. The outline of the propidium iodide counterstain was traced in FIJI, an ImageJ based analysis software and the intensity of the green channel measured using the pixel intensity measuring tool in the base software package (Schindelin *et al*., 2012; Schneider *et al*., 2012). The mean intensity of the traced area was then averaged, and the standard error calculated. Unpaired T-tests with Welch's correction were performed to determine significant difference compared to the PAF/GNA condition (p values of < 0.05 were considered significant).

Chapter 3 |Production, purification and efficacy of recombinant PA1b based fusion proteins

3.1 Introduction

This chapter focuses on attempting to produce high purity recombinant pea albumin – GNA fusion protein (PA1b/GNA) under native conditions to confirm preliminary bioassay data obtained by Dr Prashant Pyati that suggested that fusion of the pea albumin toxin peptide (PA1b), to *Galanthus nivalis* agglutinin (GNA) had substantially enhanced oral activity as compared to PA1b alone. However, Dr Pyati had encountered numerous problems in the production of high quality recombinant PA1b/GNA and preliminary bioassays were conducted using low purity fusion protein. A number of different approaches taken in an attempt to facilitate the recovery of good quality PA1b/GNA with the ultimate goal of substantiating preliminary bioassay results are reported in this chapter.

Pichia pastoris was selected as the expression organism, as previous attempts to express PA1b in *Escherichia coli* were unsuccessful (Dr Pyati, unpublished results). Furthermore, *P. pastoris* is easily genetically modified, and as a eukaryotic organism it is capable of post-translational modifications which allow the formation of disulphide bridges, of which PA1b has three (Jouvensal *et al*., 2003). It can also be grown to high cell densities using inexpensive media under highly controlled conditions, recombinant protein expression levels of over a gram per litre have been recorded in the literature and due to the *S. cerevisiae* α-factor secretory signal, the recombinant proteins are exported to the culture medium, avoiding the need for cell lysis prior to protein purification (Macauley-Patrick *et al*., 2005).

3.2 Constructs of recombinant PA1b and GNA fusion proteins: PA1b/GNA, GNA/PA1b, PA1b/EL/GNA (EL), PA1b/EL/GNA/7xH (7xH) and Strep/PA1b/EL/GNA

Gene fragments encoding PA1b and GNA proteins were codon optimised for expression in *P. pastoris* and synthesised commercially (ShineGene). The appropriate restriction sites or modifications were added by polymerase chain reaction (PCR) site directed mutagenesis. The PCR products were restricted and subsequently ligated into pGAPzαB (pGAPzαA for GNA/PA1b) in frame with the α-factor secretory signal. The initial PA1b/GNA expression construct (previously generated by Dr Pyati) and four new fusion protein constructs are described: (1) GNA/PA1b where the PA1b toxin sequence is linked to the C- (rather than N-) terminus of GNA and the histidine tag is located at the N- (rather than C-) terminus , (2) PA1b/EL/GNA that encodes a longer linker region (Gly-Gly-Gly-Ser-Ala-Ala-Ala) as compared to the original PA1b/GNA expression construct (3) PA1b/EL/GNA7xH that contains an additional C-terminal

histidine residue and (4) Strep/PA1b/EL/GNA that contains an N-terminal Strep tag sequence replacing the original C-terminal his-tag. The nucleotide and deduced amino acid sequences from sequenced constructs are shown in Figures 3.1-3.5.

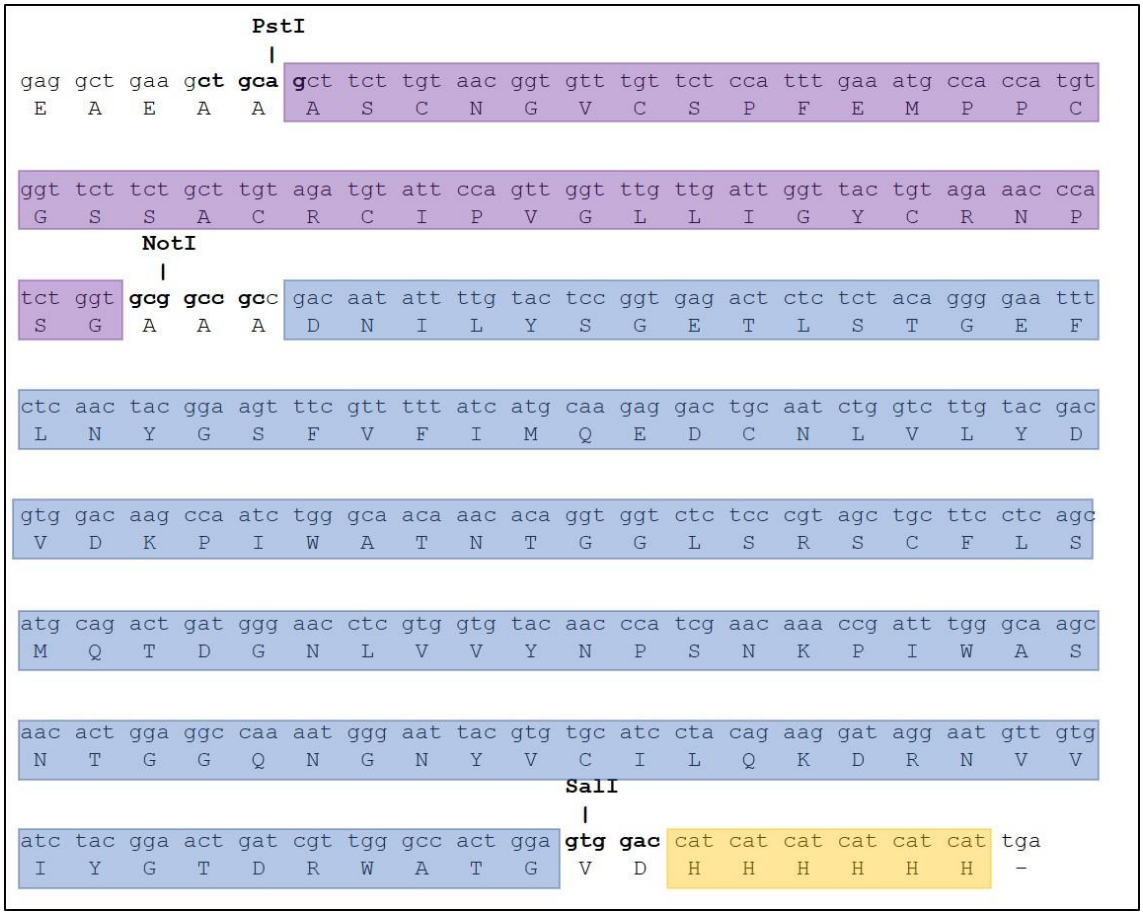


Figure 3.1 PA1b/GNA nucleotide and derived amino acid sequence from a sequenced expression construct generated by Dr Pyati. PA1b (purple), GNA (blue), his-tag (yellow). **Bold:** Restriction sites used in cloning. Predicted Mw of polypeptide: 16.5 kDa. Predicted pI: 5.78.

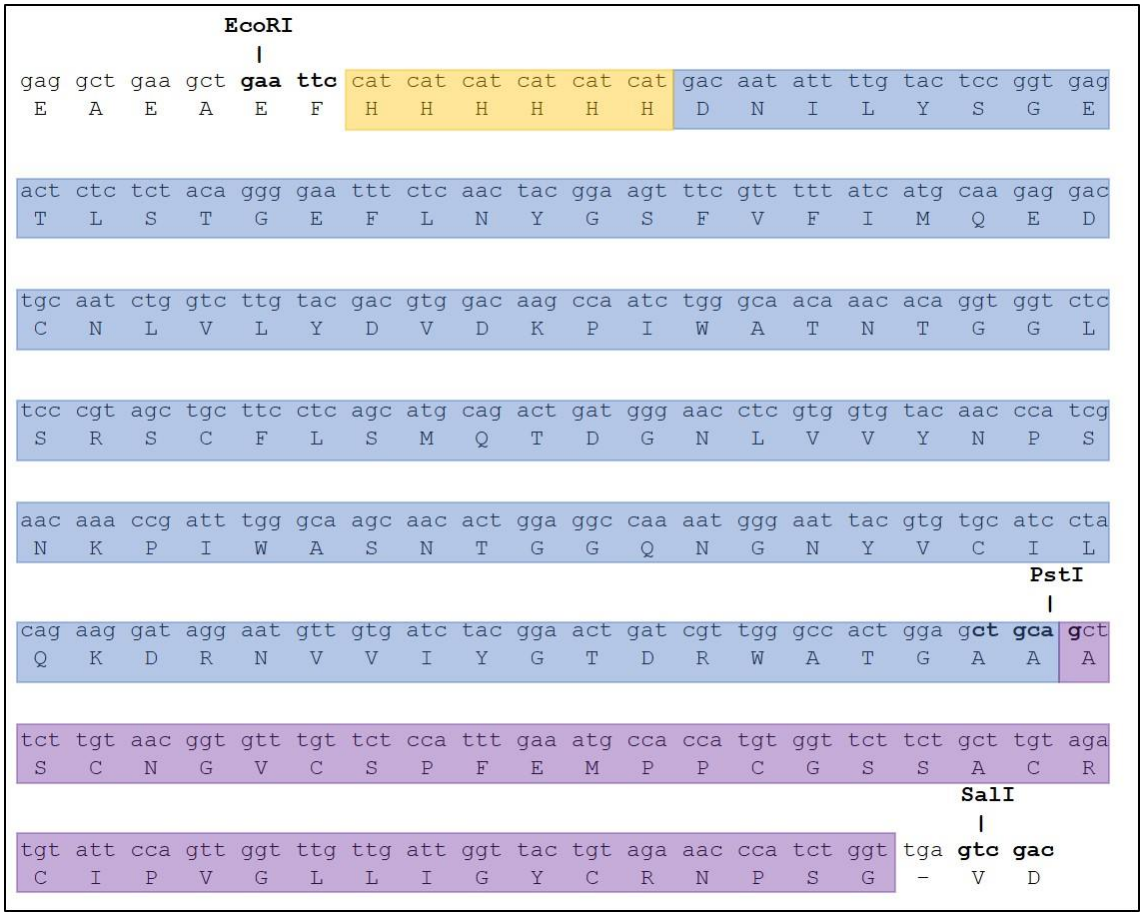


Figure 3.2 GNA/PA1b nucleotide and derived amino acid sequence, from a sequenced expression construct; PA1b (purple), GNA (blue), his-tag (yellow). **Bold:** Restriction sites used in cloning. Predicted Mw of polypeptide: 16.5 kDa. Predicted pI: 5.73.

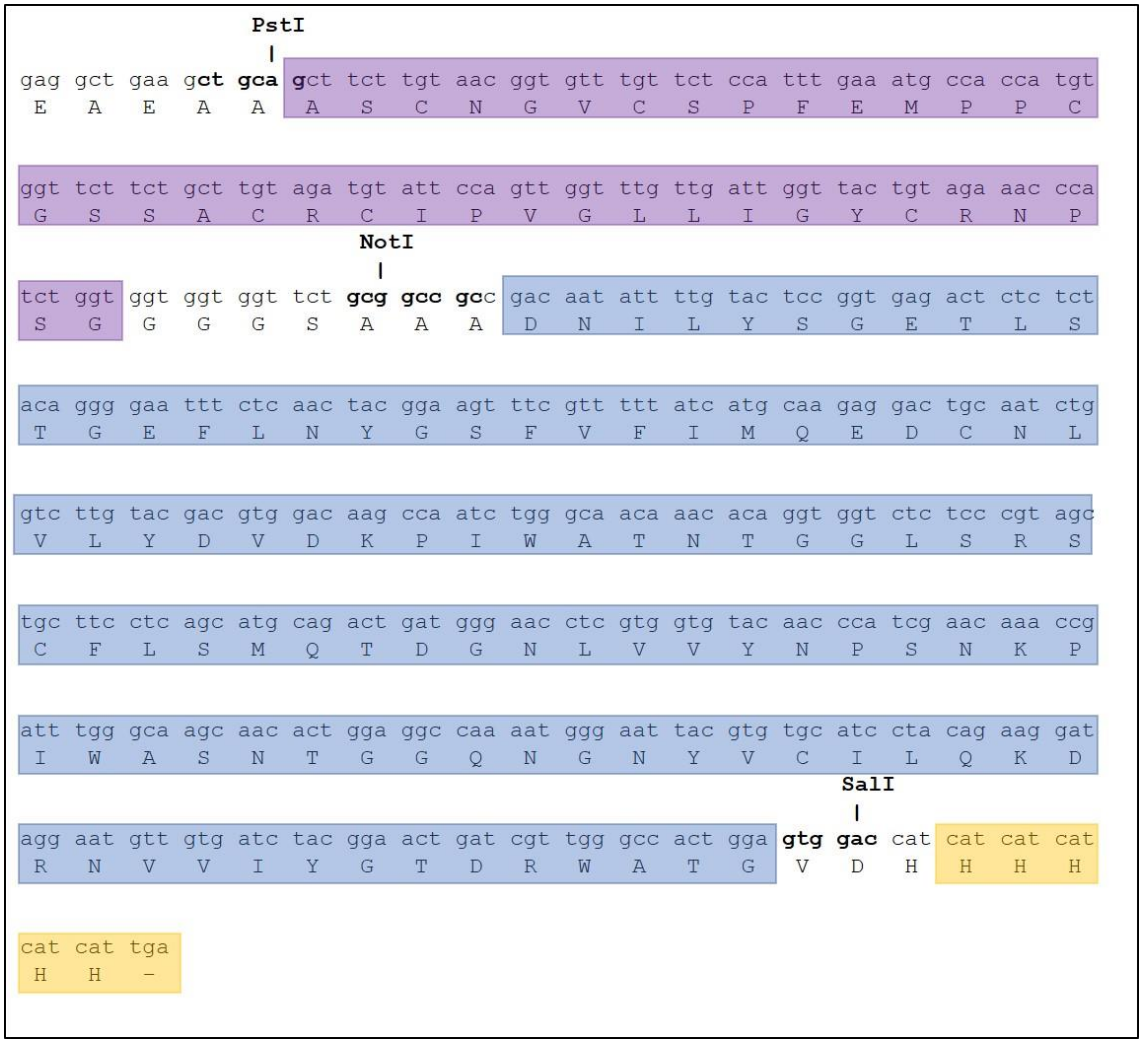


Figure 3.3 PA1b/EL/GNA nucleotide and amino acid sequence, from a sequenced expression construct. PA1b (purple), GNA (blue), his-tag (yellow). Bold: Restriction sites used in cloning. Predicted Mw of polypeptide: 17.3 kDa. Predicted pI: 5.33.

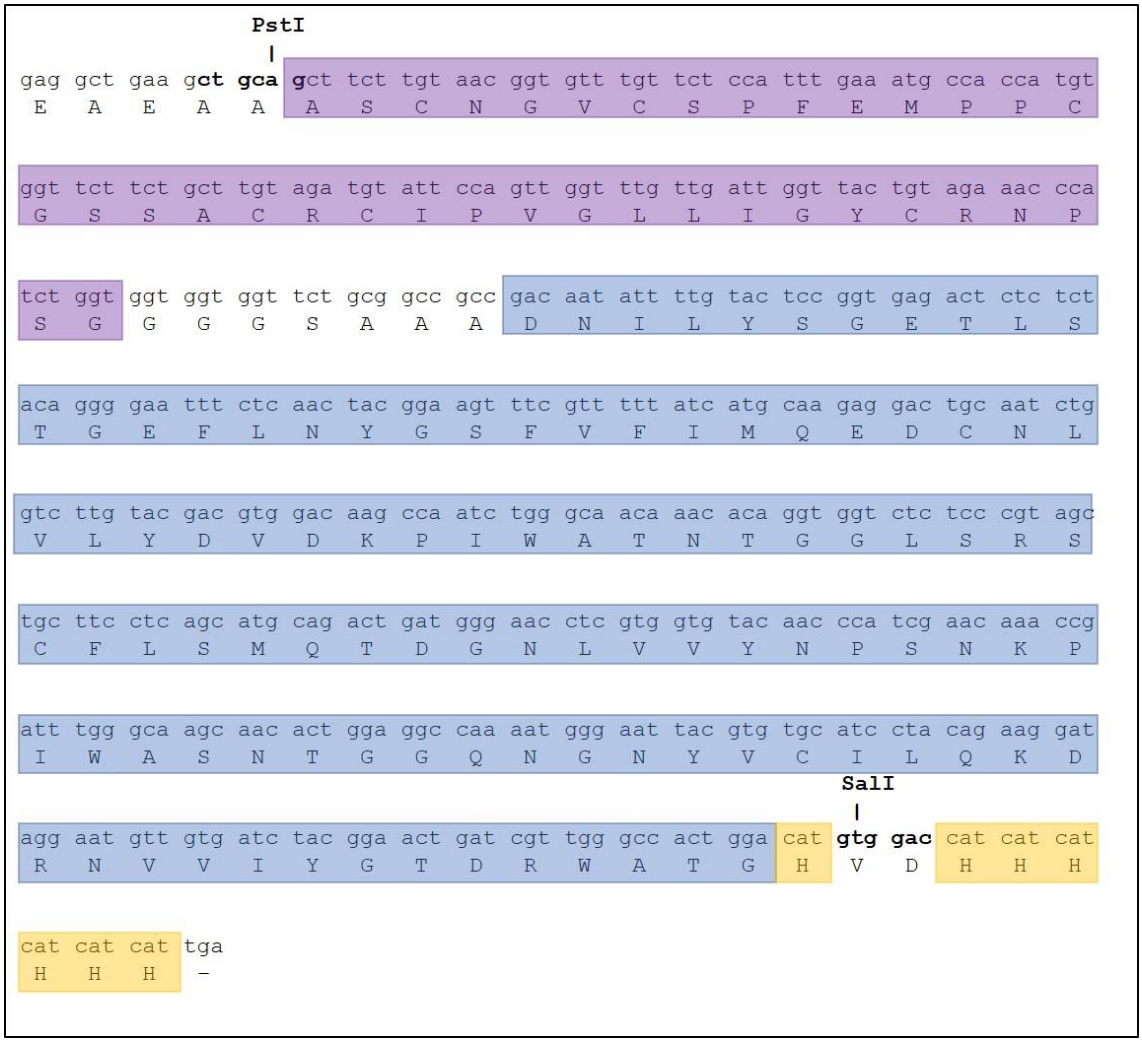


Figure 3.4 PA1b/EL/GNA7xH nucleotide and amino acid sequence, from a sequenced expression construct. PA1b (purple), GNA (blue), his-tag (yellow). Bold: Restriction sites used in cloning. Predicted Mw of polypeptide: 17.5 kDa. Predicted pI: 5.51.

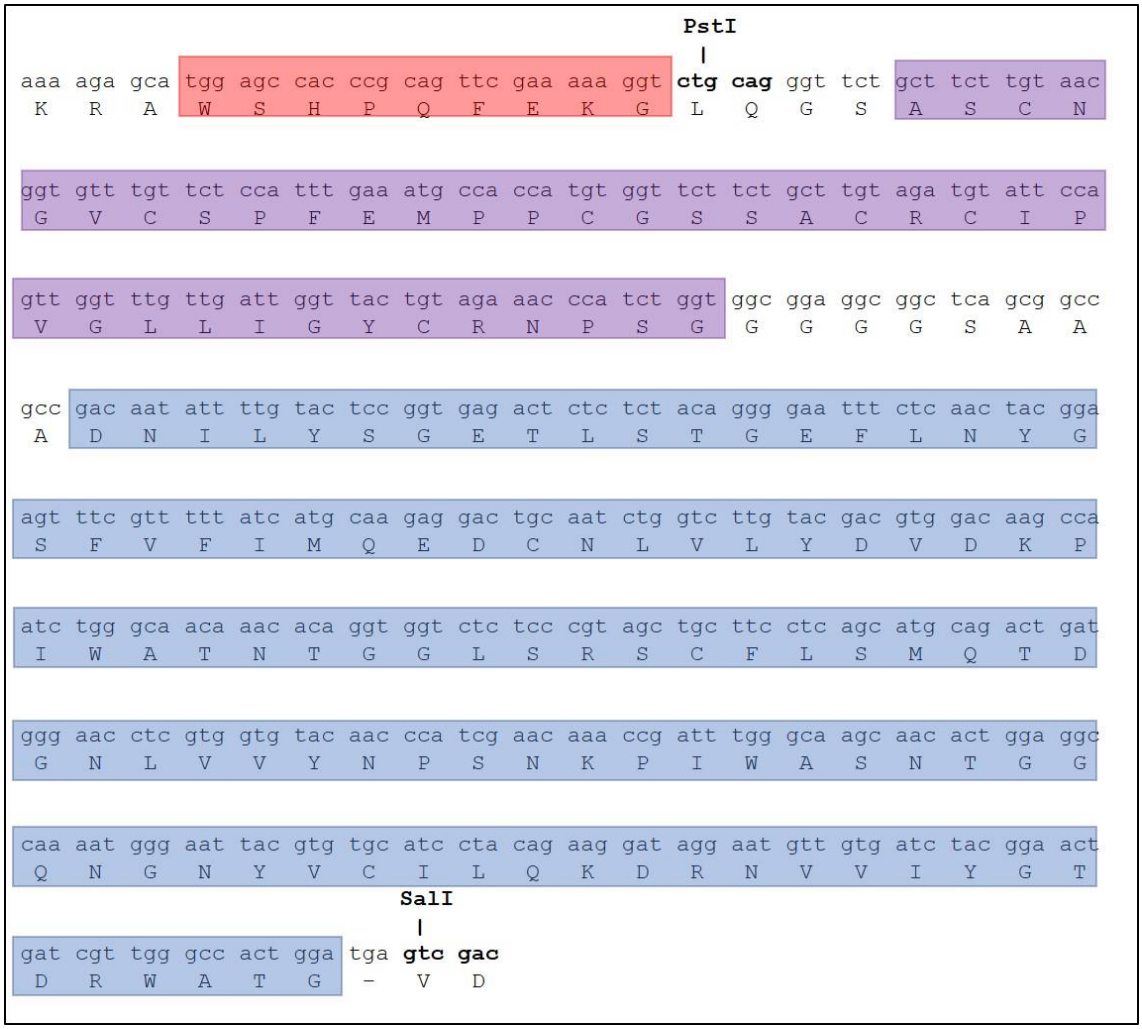


Figure 3.5 Strep/PA1b/GNA nucleotide and amino acid sequence, from a sequenced expression construct. PA1b (purple), GNA (blue), strep-tag (red). Bold: Restriction sites used in cloning. Predicted Mw of polypeptide: 17.4 kDa. Predicted pI: 4.81.

3.3 Analysis of PA1b/GNA

Prior to this project Dr Pyati had produced PA1b/GNA in *P. pastoris*, using a labour-intensive purification process (see section 2.8.2) that was only able to generate small amounts of relatively low-quality fusion protein. Lyophilised PA1b/GNA samples were a dark green sticky powder as opposed to white and fluffy as is typically observed in our laboratory. These samples did not dissolve well in PBS and pellets were observed after re-suspended samples were clarified by centrifugation in preparation for analysis by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). As shown in Figure 3.6, two major proteins corresponding to the predicted mass of 16.5 kDa for PA1b/GNA were visual in SDS-PAGE gels stained for total proteins although several contaminating proteins are also present. Both major proteins were subject to tryptic digestion and peptide fragments were analysed by LC-MS to determine if either or both contained the predicted amino acid sequence for PA1b/GNA. As depicted in Figure 3.7 B and C both proteins contained all predicted residues indicating that both were full length products. The difference in mass between the bands could be due to differential glycosylation. Differential mass on SDS-PAGE has been observed previously with the GNA fusion protein PI1a/GNA, PNGase F treatment of PI1a/GNA resulted in a single band when analysed by SDS-PAGE confirming the differential mass was due to glycosylation (Yang *et al.*, 2014). The activity of the GNA component of the fusion protein was tested for its ability to agglutinate rabbit erythrocytes, but as shown in Figure 3.8, PA1b/GNA did not display any agglutination activity. Finally, the insecticidal activity of this PA1b/GNA sample was tested against *A. pisum* in an aphid feeding assay. As shown in Figure 3.9**Error! Reference source not found.**, control survival was 90 % by day 6 of the assay whereas PA1b/GNA at a final dietary concentration of 0.05 mg mL^{-1} caused 100 % mortality after 4 days of feeding. This was in comparable to results recorded previously by Dr Pyati (unpublished). Given the lack of evidence for GNA functionality together with the presence of contaminating proteins an alternative production strategy for PA1b/GNA was sought. It was hypothesised that poor recovery from fermented culture supernatants was related to the low pH (4.5-5.5) of samples loaded onto affinity columns and that binding and hence recovery may be enhanced if samples were loaded at a pH above the isoelectric point (pI) of the protein (pI: 5.78).

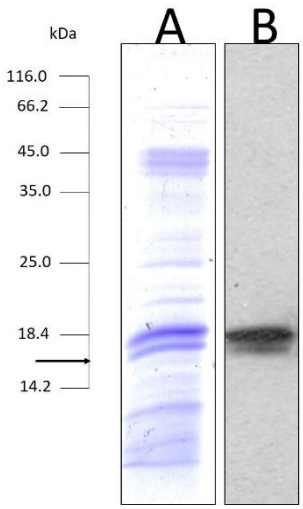


Figure 3.6 SDS-PAGE (A) and western blot (B) analysis analysis of PA1b/GNA. A: Poly acrylamide gel stained for total protein with Coomassie brilliant blue. B: Western blot of resuspended PA1b/GNA probed with anti-his antibodies. Position of protein marker mix (Thermo) run on the same gel is stated on the left-hand side. The arrow indicates the predicted protein mass.

(A) PA1b/GNA primary structure

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPPF
SNSTNNGLLFINTIASIAAKEEGVSLEKREAEAAASCNGVCSPFEMPPCGSSACRC
IPVGLLIGYCRNPSGAAADNILYSGETLSTGEFLNYGSFVFIMQEDCNLVLYDVDK
PIWATNTGLSRSCFLSMQTDGNLVVYNPSNKPIWASNTGGQNGNYVCILQKDRNVV
IYGTDRWATGVDHHHHHH

(B) PA1b/GNA 16 kDa band LC-MS/MS data

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPPF
SNSTNNGLLFINTIASIAAKEEGVSLEKREAEAAASCNGVCSPFEMPPCGSSACRC
IPVGLLIGYCRNPSGAAADNILYSGETLSTGEFLNYGSFVFIMQEDCNLVLYDVDK
PIWATNTGLSRSCFLSMQTDGNLVVYNPSNKPIWASNTGGQNGNYVCILQKDRNVV
IYGTDRWATGVDHHHHHH

(C) PA1b/GNA 15 kDa band LC-MS/MS data

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPPF
SNSTNNGLLFINTIASIAAKEEGVSLEKREAEAAASCNGVCSPFEMPPCGSSACRC
IPVGLLIGYCRNPSGAAADNILYSGETLSTGEFLNYGSFVFIMQEDCNLVLYDVDK
PIWATNTGLSRSCFLSMQTDGNLVVYNPSNKPIWASNTGGQNGNYVCILQKDRNVV
IYGTDRWATGVDHHHHHH

Figure 3.7 Primary structure and LC-MS data from the PA1b/GNA 16 kDa and 15 kDa bands on SDS-PAGE. A: Predicted primary structure of the recombinant protein expressed by transformed *P. pastoris* cells. The alpha factor secretory signal is depicted in red, PA1b in purple, PA1a in green, GNA in blue and the his-tag in yellow. Remaining residues are the result of the cloning process for the expression construct. B and C show the detected peptides from both bands. Colours denote the confidence with which each peptide sequence was matched to the expected sequence: Black = not detected, Green = high, Yellow = medium, Red = poor.

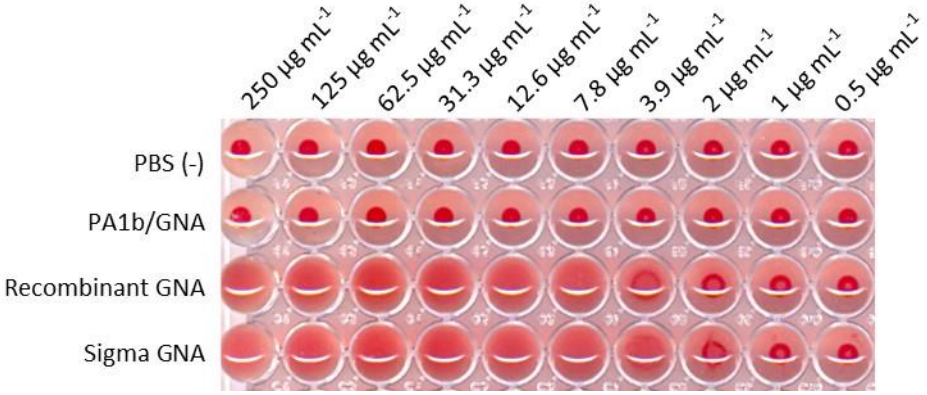


Figure 3.8 Agglutination assay to detect function of the GNA subunit of PA1b/GNA. Protein was serially diluted and incubated with a 2 % (v/v) rabbit erythrocyte solution. After 2 hours the plate was imaged to determine the lowest concentration at which GNA can crosslink rabbit erythrocytes. PBS was used as a negative control and recombinant GNA and Commercial (Sigma-Aldrich, St. Louis, USA) GNA were included as positive controls. Agglutinated or crosslinked cells display an even red colour in the well whilst non-agglutinated cells settle to the bottom of the plate and form a visible dark red pellet.

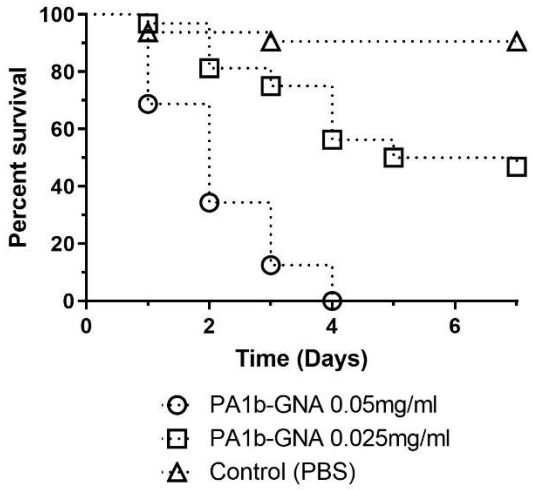


Figure 3.9 Survival of *A. pisum* fed on artificial diet containing recombinant PA1b/GNA. Purified recombinant protein re-suspended in PBS was incorporated into artificial diet to give final PA1b/GNA concentrations of 0.05 mg mL^{-1}, and 0.025 mg mL^{-1}, the control treatment consisted of the diet with an equivalent volume of PBS to that present in the protein treatments. N=2.

3.4 Expression and purification of PA1b/GNA from different fermentation media

The PA1b/GNA yeast expression clone created by Dr Pyati was used in benchtop fermentation experiments using different media, namely basal salt (BSM), yeast peptone glycerol medium (YPG), and buffered Glycerol-complex medium (BMGY) at pH 4.5, pH 6.5 and pH 6, respectively (sections 2.6). Culture pH was maintained by the addition of 17.5% ammonia and glycerol was used as a carbon source in all fermentations. As a result of the α-factor secretory signal all protein was secreted into the growth medium. Expression level was estimated by comparison of culture supernatant and known protein standards, by western blotting for all three fermentation media using anti-GNA antibodies.

The results of the expression analysis are shown in Figure 3.10. Two immunoreactive proteins were present at approximately 15 kDa and 16 kDa in all supernatants when analysed by western blots probed with anti-GNA antibodies. The predicted mass of PA1b/GNA is 16.5 kDa. The final expression level of PA1b/GNA in BSM was estimated at 40 mg L^{-1} and a lower 20 mg L^{-1} in YPG, but the protein content seems to have peaked at 48 hours post inoculation, and as there is no evidence for GNA-immunoreactive degradation products, this suggests that the cells stopped replicating, although wet pellet weights were not recorded. In BMGY the expression level was estimated at 25 mg L^{-1}.

Clarified culture supernatants were obtained by centrifugation and filtration of fermented cultures, and proteins were then loaded onto nickel affinity chromatography columns (Section 2.8.1). Purification fractions were analysed to determine if the use of different fermentation media with different pHs influenced the binding efficiency of PA1b/GNA to nickel matrices. PA1b/GNA fermented in BSM was purified following the purification protocol developed by Dr Pyati, section 2.8.2, as previous work shows purification following the standard protocol was not viable and resulted in precipitation of the purified protein during dialysis. After initial purification under native conditions the remaining protein was purified under denaturing conditions, as it was noted that its binding efficiency to the column was poor and much protein was lost during washing under native conditions, the addition of 6 M urea to the culture supernatant remedied this issue (Figure 3.11). The remaining two fermentations were purified following the standard protocol (section 2.8.1). No protein was visible in the flow through from the BMGY supernatant, and protein was eluted in the 200 mM imidazole fraction (Figure 3.11). Approximately 50 % of the protein visible in the YPG supernatant bound to the column under native conditions; and no protein was lost in the washing steps unlike with BSM. Protein present in the BMGY appeared to bind the most efficiently to the nickel columns under native conditions.

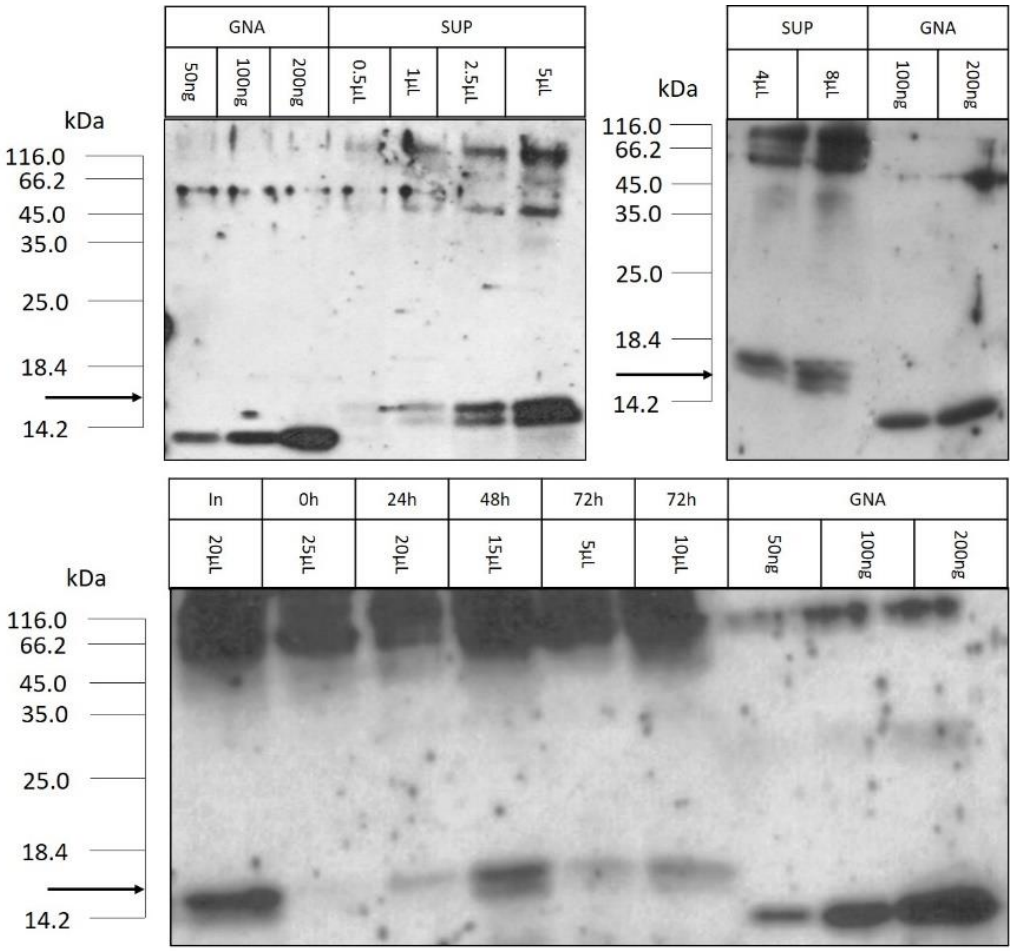


Figure 3.10 Expression levels of PA1b/GNA fermented in BSM (upper left), BMGY (upper right) and YPG (bottom). Western blot probed with anti-GNA antibodies of fermented culture supernatant (SUP) expressing PA1b/GNA. GNA: recombinant GNA standards loaded at the quantity stated. The culture supernatant was loaded in each land at the volume stated. In blot C lane labels denote the time at which that sample was removed from the fermenter. Position of protein marker mix (Thermo) run on the same gel is stated on the left-hand side. The arrow indicates the expected protein mass.

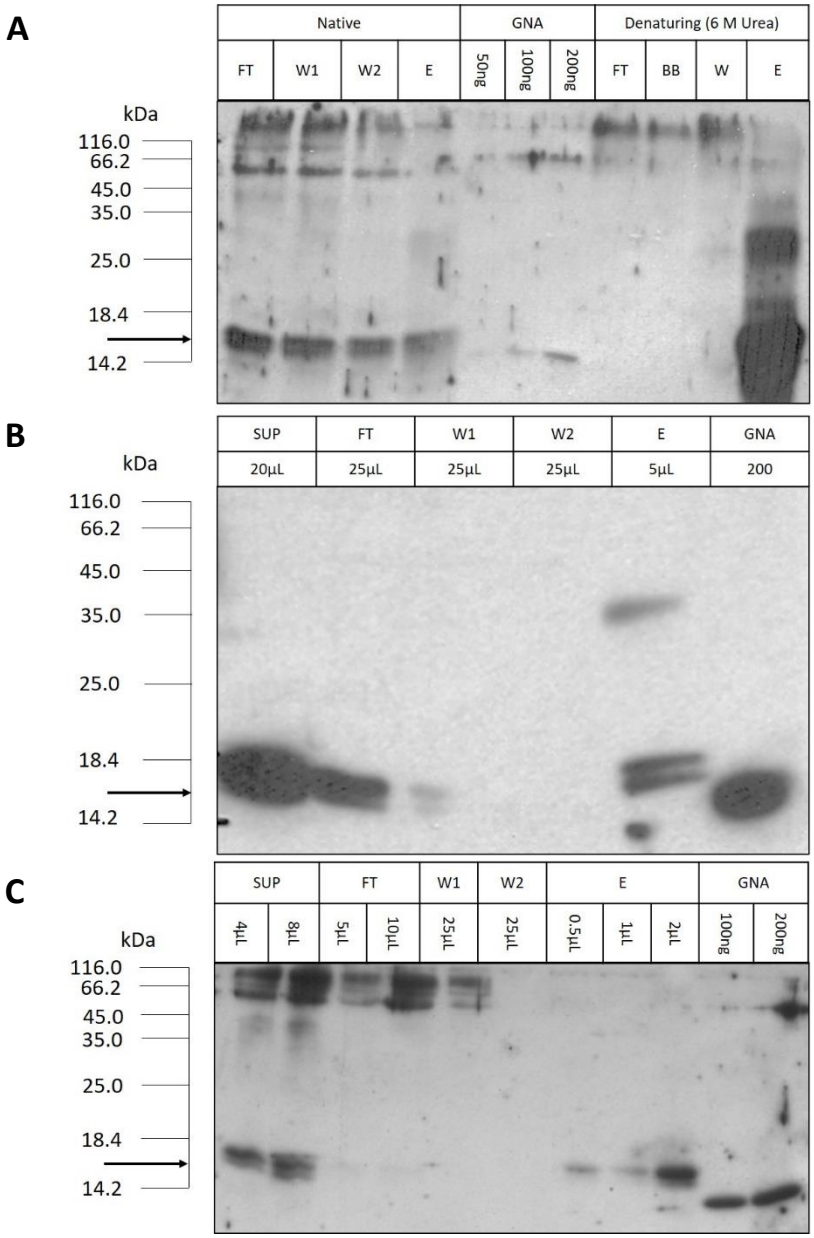


Figure 3.11 Purification profiles of PA1b/GNA from BSM (A), YPG (B) and BMGY (C). Western blots were probed with anti-GNA antibodies. A: Comparison of fractions from fermentation using BSM under native and denaturing conditions, 20 µL was loaded per lane for native purification, and equivalent volumes were loaded for denaturing purification. B: Purification fractions of PA1b/GNA purified from fermentation using YPG medium. C: Purification fractions from fermentation using BMGY medium. SUP: Supernatant, FT: flow-through, W1: binding buffer, W2: binding buffer + 10 mM imidazole, E: binding buffer + 200 mM imidazole. GNA: recombinant GNA standards quantity stated above the lane. Volumes of fraction loaded is stated above the lane for blot B and C. Position of protein marker mix (Thermo) run on the same gel is stated on the left-hand side. The arrow indicates the expected protein mass.

The PA1b/GNA containing peak fractions were collected, de-salted by dialysis against water at 4 °C overnight, and freeze-dried. In the case of the BSM fermentation the peak fractions eluted under denaturing conditions were dialysed against a decreasing urea concentration gradient. The ratio of fusion protein to lyophilised powder was estimated semi-quantitatively by comparison on SDS-PAGE with commercial GNA standards, the quantification results are shown in Figure 3.12. There was no soluble PA1b/GNA present in the lyophilised powder derived from peaks eluted under denaturing conditions (and using BSM fermentation medium). However, protein was detected when the pellet from PBS resuspension/centrifugation was resuspended in SDS loading buffer and loaded on SDS-PAGE suggesting that the little protein that was present failed to re-fold during dialysis. In the lyophilised powder derived from the YPG fermentation the proportion of PA1b/GNA was 2 % (w/w); when stained for total protein a second major band at approximately 40 kDa is visible and there was background visible in the lane suggesting contamination. The estimated proportion of PA1b/GNA in the lyophilised powder from the BMGY fermentation was 15 % (w/w), however again many other proteins were present. The total protein recovered from the YPG and BMGY fermentation runs was 0.6 mg and 5 mg respectively, resulting in a recovery of 1.1 % and 8 % of total available protein in fermented supernatants.

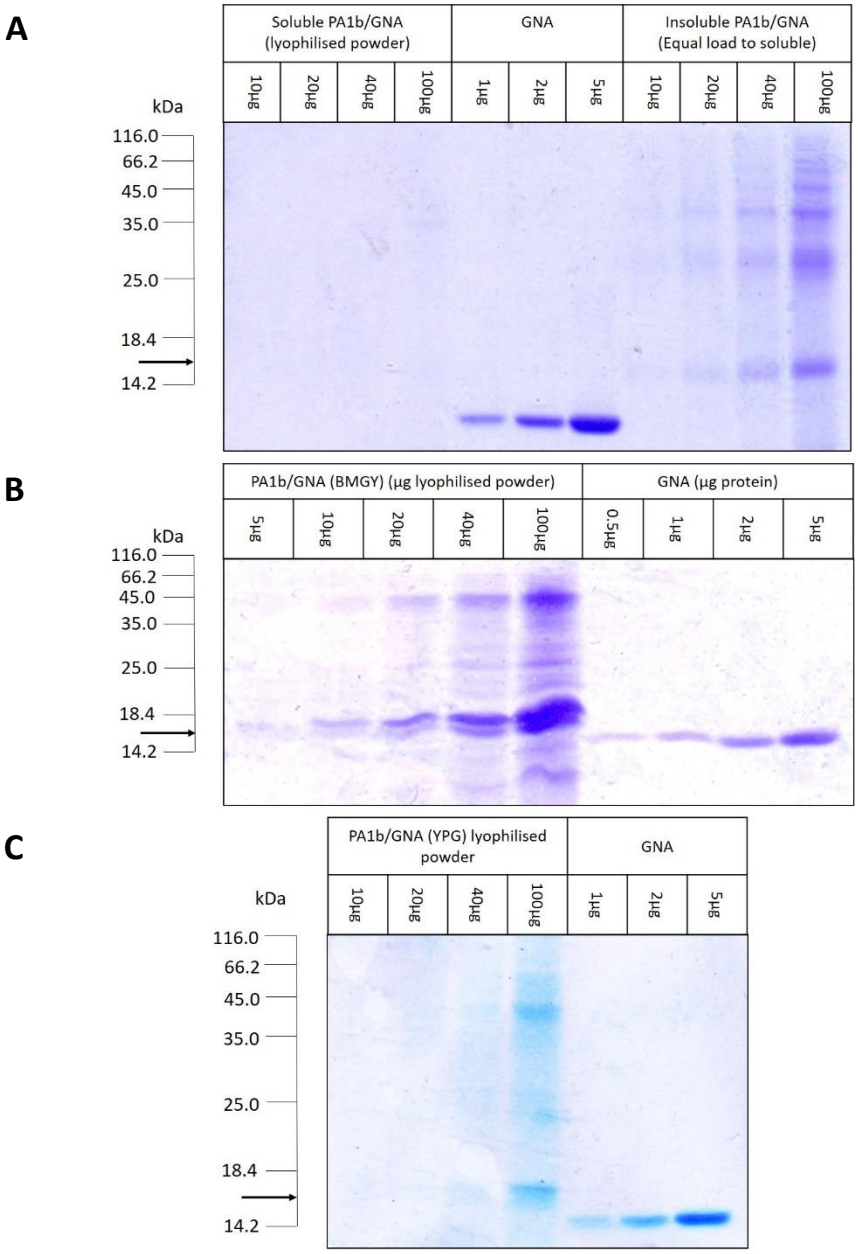


Figure 3.12 Quantification of purified and lyophilised PA1b/GNA samples fermented in BSM (A), BMGY (B) and YPG (C). Coomassie brilliant blue stained SDS-PAGE gels. Quantity of lyophilised powder and commercial GNA standards are stated above the lanes in micrograms. Position of protein marker mix (Thermo) run on the same gel is stated on the left-hand side. The arrow indicates the expected protein mass.

3.5 Expression and purification of GNA/PA1b, PA1b/EL/GNA, PA1b/EL/GNA7xH and Strep/PA1b/GNA

Sequence confirmed plasmid DNA for each construct was linearized by BlnI and transformed into protease deficient *P. pastoris* SMD1168H cells, and transformants were selected on zeocin containing plates. Colonies were screened for the expression of recombinant proteins by growing in 10 mL YPG shake flasks for 48 – 72 hours at 30 °C and western blot analysis (anti-GNA or anti-His antibodies) of culture supernatant samples. Results of one of the expression screens are shown in Figure 3.13.

No immunoreactive bands were present in any of 40 clones screened for the expression of GNA/PA1b, despite screening of colonies obtained from two different yeast transformations (Figure 3.13, example of one of the screens). Integration of the recombinant gene in to the *P. pastoris* genome was confirmed by yeast genomic DNA extraction followed by PCR using expression vector-specific primers. PCR products were detected at the expected mass 870 bp, 470 bp fusion protein + 400 bp from vector-specific primers) in two of the four randomly selected transformants Figure 3.14, Section 2.5). This suggests that, despite successful integration of the expression cassette into the yeast genome, the yeast cells were unable to express the fusion protein when the toxin was linked to the C-terminus of GNA. By contrast, immunoreactive bands of approximately 18 kDa (close to predicted masses) were detected in a small number of screened PA1b/EL/GNA and PA1b/EL/GNA7xH transformants (Figure 3.15 A and B). Expression screening of Strep/PA1b/GNA transformants detected only one transformant expressing an immunoreactive protein at approximately 16 kDa, 1.5 kDa below the expected mass of this protein (Figure 3.15C). PA1b/EL/GNA and Strep/PA1b/GNA both appeared as single bands on western blots suggesting minimal cleavage of the fusion protein by yeast proteases. PA1b/EL/GNA7xH clones showed the presence of two immunoreactive bands, albeit of low intensity, of similar to predicted masses in clone 4 and clone 10. The presence of the purification tag was confirmed using tag specific antibodies (not shown). The best expressing clones were selected for benchtop fermentation which was conducted in a 7.5 L bioreactor following set environmental conditions in BSM (2.6).

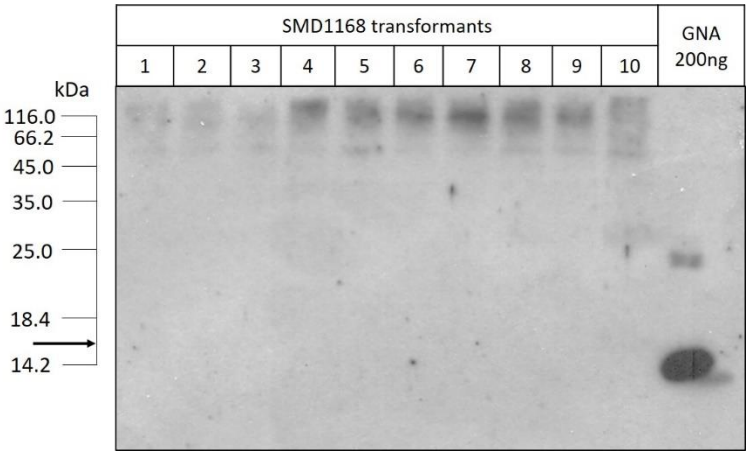


Figure 3.13 Screening of small-scale yeast cultures for GNA/PA1b expression. Western blot (anti-GNA antibodies) screening to select a GNA/PA1b expressing *P. pastoris* SMD1168 clone. Individual colonies were inoculated into 10 mL YPG flasks with 100 µg mL^{-1} zeocin, incubated for 72h at 30 °C with shaking. Cultures were centrifuged and individual culture supernatant (25 µL) was loaded in each lane (1-10). GNA 200 ng was included as a positive control. Position of protein marker mix (Thermo) run on the same gel is stated on the left-hand side. The arrow indicates the expected protein mass.

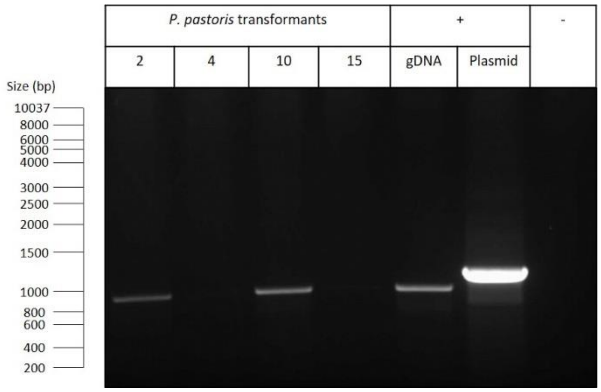


Figure 3.14 Yeast colony PCR to confirm the transformation of *P. pastoris* colonies. The numbers represent randomly selected colonies which had previously been found not to express GNA/PA1b but grew well in selective media. + depicts the positive controls, gDNA was a yeast clone known to express PA1b/GNA, plasmid was His/GNA/Omega/pGAPzαB, a negative control was PCR reaction without yeast gDNA. All reactions used the same primers Forward: pGAP fwd, Reverse: AOX 3' rev. Position of DNA marker mix, Hyperladder 1 kb (Bioline) run on the same gel is stated on the left-hand side.

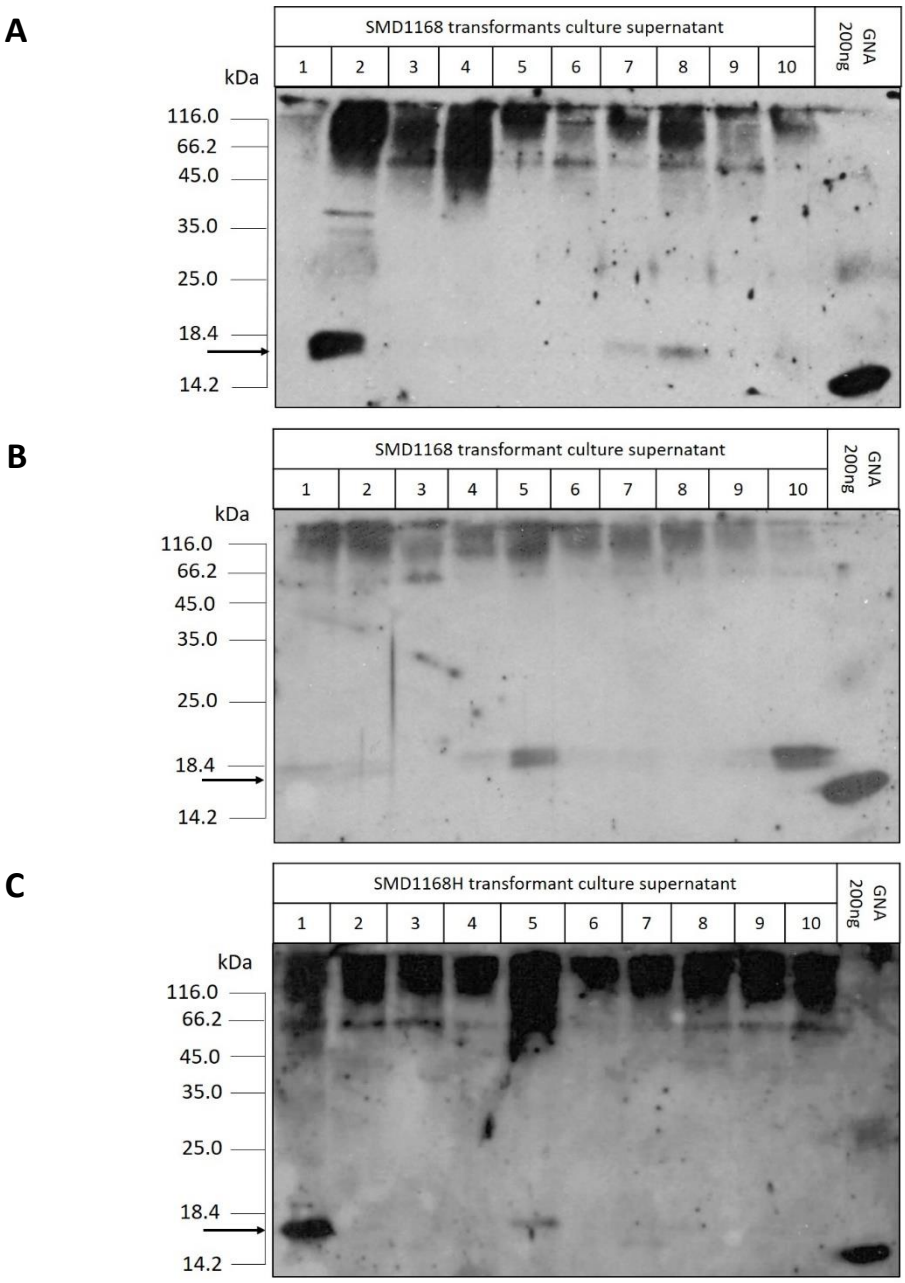


Figure 3.15 Western blot screening (anti-GNA antibodies) of small-scale *P. pastoris* SMD1168 cultures for the expression of A: PA1b/EL/GNA, B: PA1b/EL/GNA 7xH, C: Strep/PA1b/GNA. Individual colonies were inoculated into 10 mL YPG flasks with 100 µg mL^{-1} zeocin, incubated for 72h at 30 °C with shaking. Cultures were centrifuged and individual culture supernatant (25 µL) was loaded in each lane (1-10). GNA 200 ng was included as a positive control. Position of protein marker mix (Thermo) run on the same gel is stated on the left-hand side. The arrow indicates the expected protein mass.

Wet pellet weights were monitored every 24 hours throughout fermentations and in all three constructs cell density increased throughout the fermentation (Table 3.1). Expression levels throughout the fermentation runs were estimated semi-quantitatively using western blotting as described previously. The results are displayed in Figure 3.16. PA1b/EL/GNA showed two immunoreactive bands at approximately its predicted molecular weight, contrary to the single band detected during the expression screening and its final expression level was estimated to be 30 mg L^{-1} culture supernatant. The expression level of PA1b/EL/GNA 7xH was estimated to be approximately 40 mg L^{-1} based on a single immunoreactive band, which may actually be two overlapping bands as seen in the screening blot Figure 3.15A. Finally, for Strep/PA1b/GNA a single immunoreactive band was present in the supernatant, and the final expression level was estimated to be 20 mg L^{-1} culture supernatant. The two bands observed in PA1b/EL/GNA when fermented contrast with what was observed in the expression screening Figure 3.15, suggesting processing akin to what was observed in PA1b/GNA for this protein construct.

Table 3.1 Wet pellet weights recorded for different PA1b fusion proteins grown in basal salt medium.

	PA1b/EL/GNA	7xH	Strep/PA1b/GNA
Time (Hours)	Wet Pellet weight (mg)	Wet Pellet weight (mg)	Wet Pellet weight (mg)
Inoculant	78.0	63.6	39.1
0	22.7	31.6	12.8
24	165.3	98.1	104.8
48	280.0	287.0	355.6
72	344.8	350.0	367.1

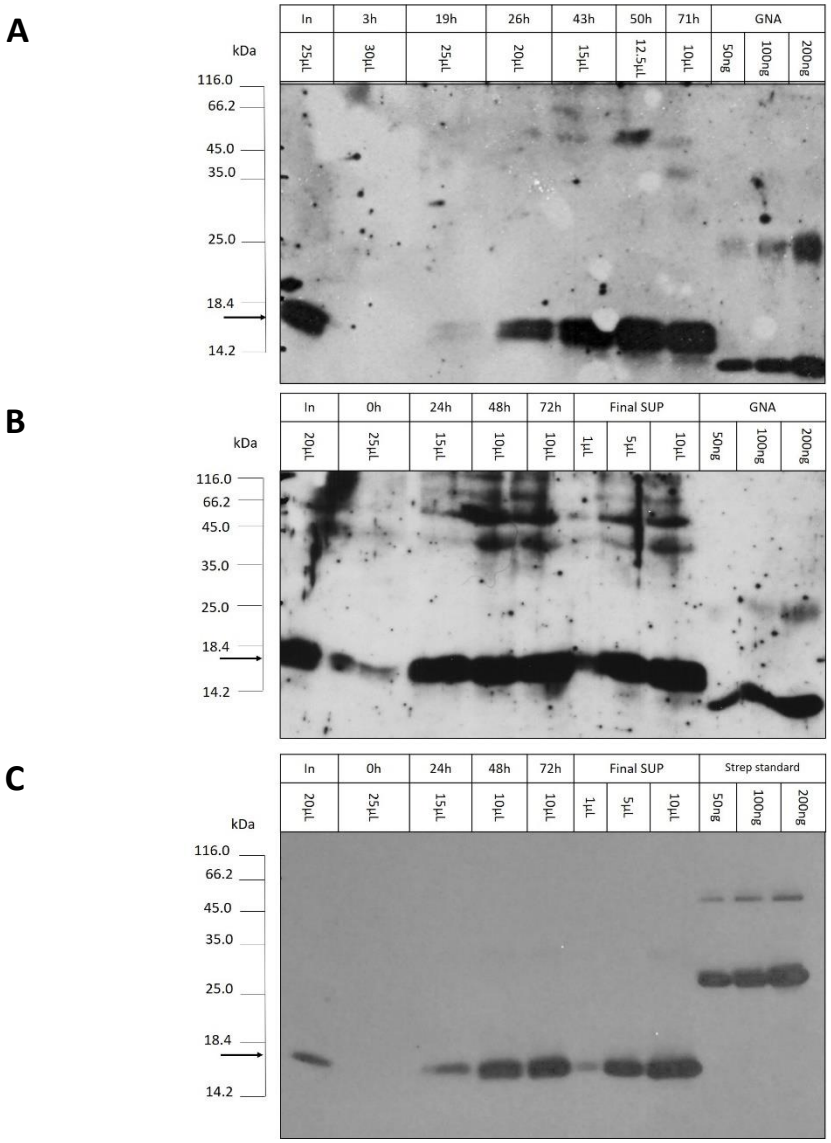


Figure 3.16 Expression levels of PA1b/EL/GNA (A), 7xH (B) and Strep/PA1b/GNA (C). Western blot probed with anti-GNA (A and B) and anti-Strep (C) antibodies of fermented culture supernatant expressing PA1b/EL/GNA (A), 7xH (B) and Strep/PA1b/GNA (C). GNA: recombinant GNA standards loaded at the quantity stated. The culture supernatant was loaded in each land at the volume stated. Position of protein marker mix (Thermo) run on the same gel is stated on the left-hand side. The arrow indicates the expected protein mass.

With the exception of Strep/PA1b/GNA which was purified using a strep-tactin affinity chromatography (section 2.9) all proteins were purified using nickel affinity chromatography. The different purification fractions were collected and analysed by western blot with anti-GNA antibodies or SDS-PAGE to assess column binding efficiency and these results are shown in Figure 3.17**Error! Reference source not found.**. PA1b/EL/GNA showed little difference in intensity between the immunoreactive band in the supernatant and the flow through, a similar pattern was observed for PA1b/EL/GNA 7xH. To test if this was due to the availability of the tag for binding to nickel a small portion of supernatant from the PA1b/EL/GNA7xH fermentation was purified under denaturing conditions (addition of 6 M urea to all buffers) and purification samples were again analysed by western blotting. As shown in Figure 3.18 a small decrease in immunoreactivity of the fusion protein was observed in the flow through and wash samples collected during purification under denaturing conditions as compared to equivalent samples collected during native purification. In addition, a small increase in immunoreactivity of the fusion protein is observed in the elution sample collected under denaturing conditions as compared to the equivalent native elution peak fraction (Figure 3.18). This suggests that even when the fusion protein was unfolded (i.e. denatured) and an additional histidine residue was incorporated into the final protein product that the interaction with nickel matrices was poor. Poor recovery from culture supernatants is thus more likely to be attributable to relatively low pH of the loaded samples preventing strong binding of the recombinant protein to the nickel matrix.

Elution fractions collected from Strep/PA1b/GNA purification when analysed by SDS-PAGE, showed the presence of a faint protein band at the expected mass of 17.3 kDa (Figure 3.17). However, as for PA1b/GNA, very little protein appears to have bound to the column as the 17.3 kDa protein is present in similar quantities in loaded and flow through fractions and is similar abundance to the loaded sample. No visible protein was lost in the washing steps, and protein was detected in both elution 1 and 2 suggesting some of the expressed protein was binding to the column (Figure 3.17). Optimisation of this purification protocol suggests that poor column binding is unlikely due to pH (data not shown), therefore is likely due to tag availability, as with PA1b/GNA and PA1b/GNA 7xH. However, there was no possibility to attempt strep-tactin purification of Strep/PA1b/GNA under denaturing conditions as the purification strategy is based on protein-protein interactions. The peak fractions were desalted then lyophilised.

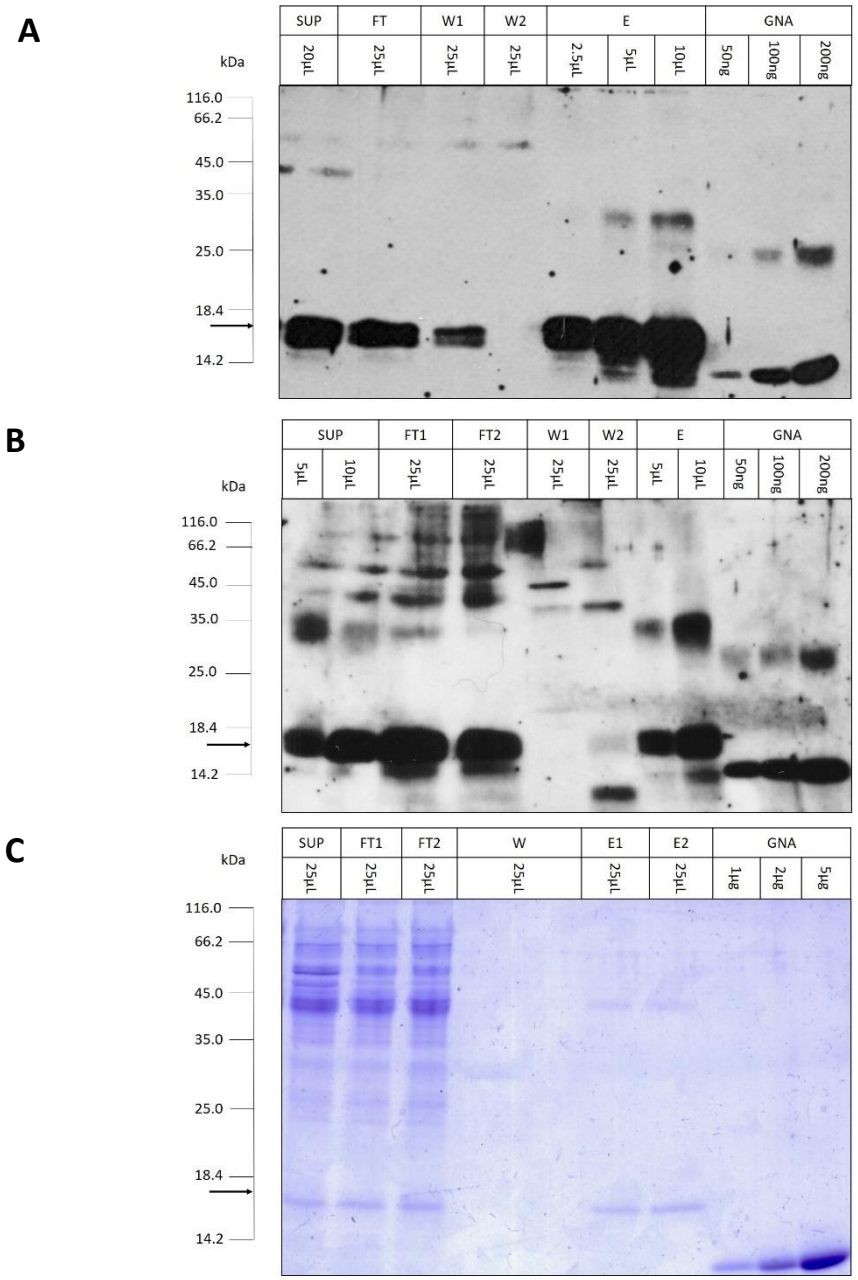


Figure 3.17 Purification fractions of PA1b/EL/GNA (A), PA1b/EL/GNA7xH (B) and Strep/PA1b/GNA (C). Western blots (A and B) probed with anti-GNA antibodies; Coomassie stain (C) of fermented culture supernatant expressing Strep/PA1b/GNA; SUP: Supernatant, FT: Flow through, W1: 1 x binding buffer, W2: 1 x binding buffer + 10 mM imidazole, E:1 x binding buffer + 200 mM imidazole, GNA: recombinant GNA standards (quantity in ng stated above lanes). Position of protein marker mix (Thermo) run on the same gel is stated on the left-hand side. The arrow indicates the expected protein mass.

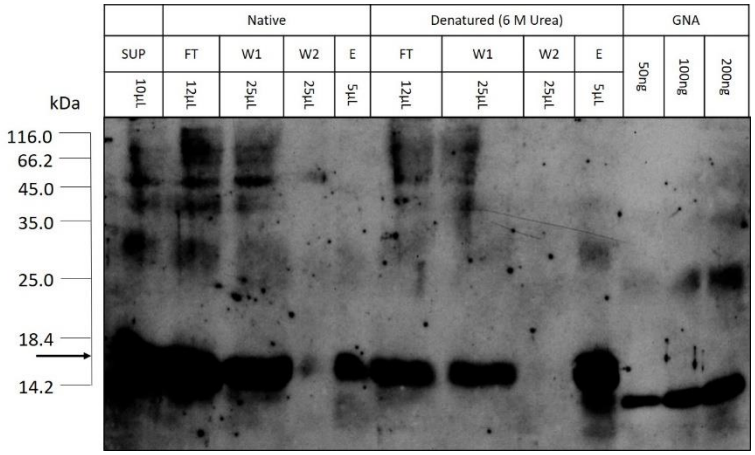


Figure 3.18 Comparison of purification fractions of PA1b/EL/GNA7xH under native and denaturing conditions. Western blot probed with anti-GNA antibodies. Numbers depict volume of supernatant loaded in microlitres. SUP: Supernatant, FT: Flow through, W1: 1 x BB + 10 mM imidazole, W2 1 x BB + 200 mM imidazole. GNA: recombinant GNA standards, quantity in ng stated above the lane. Position of protein marker mix (Thermo) run on the same gel is stated on the left-hand side. The arrow indicates the expected protein mass.

The protein content in lyophilised samples was estimated by comparison of staining intensity of recombinant proteins (of known powder weight) on Coomassie stained gels with commercial GNA standards of known concentration. As shown in Figure 3.19, lyophilised PA1b/EL/GNA and PA1b/EL/GNA7xH samples contained relatively high amounts of contaminating proteins in gels stained for total protein and the amount of fusion protein present (considering both 17.3 and 17.5 kDa proteins shown to be immunoreactive with anti-His and anti-GNA antibodies) was estimated to be 6 % and 2 % (w/w), respectively. The total protein recovered from 3 L of supernatant of PA1b/EL/GNA and PA1b/EL/GNA 7xH was 3.4 mg and 0.7 mg, respectively. Strep/PA1b/GNA was also estimated to contain 2 % (w/w) total protein, but this sample was relatively pure containing only 1 major contaminating protein (approximately 45 kDa) in the lyophilised powder; 1 mg of Strep/PA1b/GNA was recovered from 3 L of culture supernatant after purification. The estimated recovery of each protein from culture supernatants was 3.7 %, 0.5 % and 1.6 % for PA1b/EL/GNA, PA1b/EL/GNA 7xH and Strep/PA1b/GNA respectively. Recovery was considerably lower than for PA1b/GNA produced in BMGY (8 %), but in the case of Strep/PA1b/GNA the protein recovered had greater purity.

3.6 Biological activity of Strep/PA1b/GNA

In order to elucidate the sequence of the two proteins present in the purified Strep/PA1b/GNA fraction the 45 kDa and 17 kDa bands were excised from an SDS-PAGE gel, trypsin digested, and analysed by LC-MS. The 17 kDa molecular band was identified as full length Strep/PA1b/GNA with a correctly processed N-terminus and an N-terminal Strep tag sequence (Figure 3.20). The high molecular weight band contained a complex mixture of *P. pastoris* proteins with the major protein identified as a cell wall protein with similarities to glucanases (accession number: C4QVL7_KOMPG), from *Komagataella spp.* (syn *Pichia spp.*). No similar sequences to the strep-tag II (WSHPQFEK) were observed in the identified protein.

Strep/PA1b/GNA was fed to *A. pisum* to determine if its insecticidal activity was comparable to that observed for PA1b/GNA. Strep/PA1b/GNA was fed at 7 different dietary concentrations between 0.4 and 0.006 mg mL^{-1}. The range of concentrations selected were chosen as they encompassed the higher concentrations for which toxicity was observed for PAF/GNA (Chapter 6) to the lowest concentrations which toxicity was observed for PA1b/GNA. As shown in Figure 3.21, Control survival 93 % on day 7. Complete mortality was observed for Strep/PA1b/GNA fed aphids at concentrations of 0.4 mg mL^{-1} and 0.2 mg mL^{-1} on day 3 and 4, respectively. Lower concentrations of Strep/PA1b/GNA resulted in reduced survival as compared to the control but did not achieve complete mortality by day 7. The LC_{50} for

Strep/PA1b/GNA was calculated at 0.1169 mg mL^{-1}. Only the two highest doses, 0.4 and 0.2 mg mL^{-1} were found to be significantly different from the control ($p < 0.0001$; log-rank).

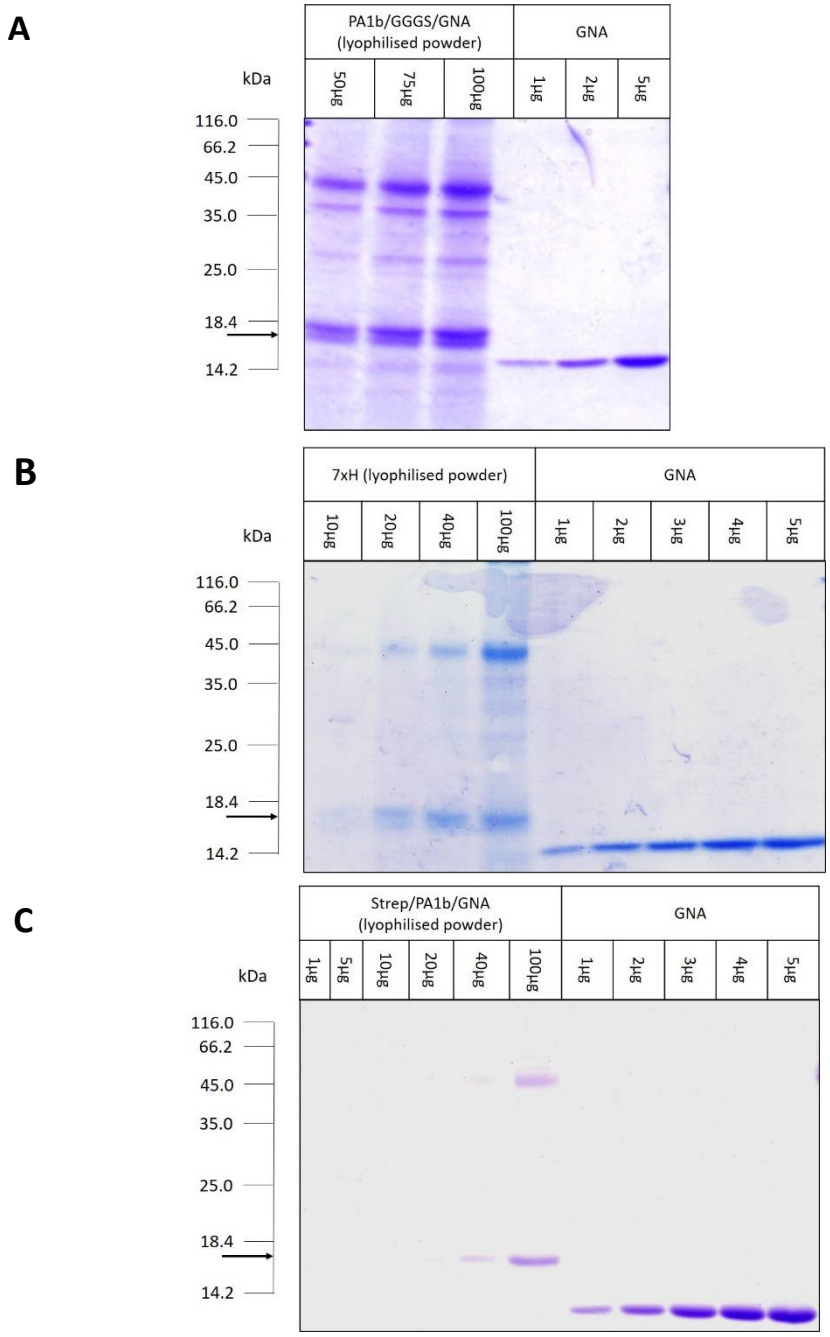


Figure 3.19 Quantification of PA1b/GGGS/GNA (A), 7xH (B) and Strep/PA1b/GNA (C). Coomassie brilliant blue stained SDS-PAGE gels (17.5% acrylamide) of recombinant PA1b/GNA after purification by nickel or streptactin affinity chromatography, lyophilisation and resuspension in PBS. Commercial GNA (Sigma-Aldrich, St. Louis, USA) was used as a known standard for comparison. Position of protein marker mix (Thermo) run on the same gel is stated on the left-hand side. The arrow indicates the expected protein mass.

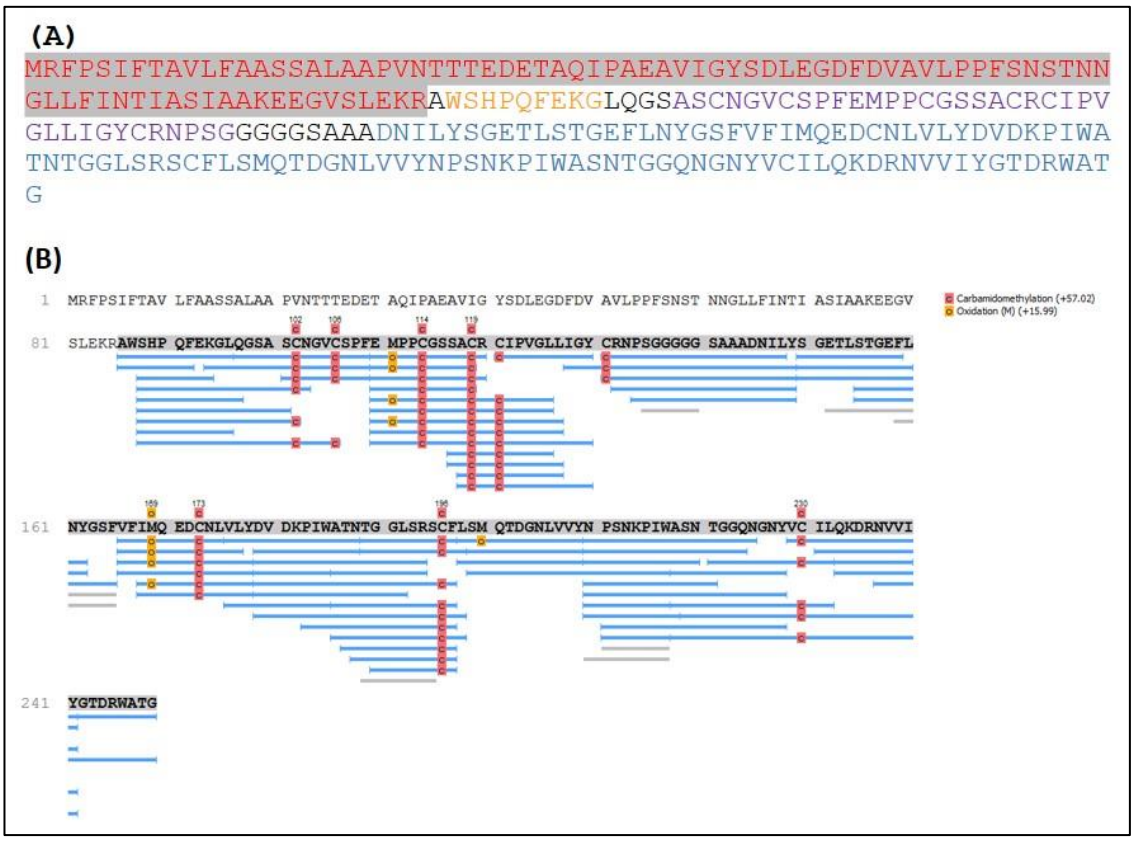


Figure 3.20 Primary structure and LC-MS data from the Strep/PA1b/GNA 17 kDa band on SDS-PAGE. A: Primary structure of the recombinant protein expressed by transformed *P. pastoris* cells. The alpha factor secretory signal is depicted in red, PA1b in purple, GNA in blue and the strep-tag in yellow. Remaining residues are the result of the cloning process for the expression construct. Residues highlighted in grey were not detected in the 17 kDa band see (B). B: Peptide matches of the 17 kDa Strep/PA1b/GNA protein.

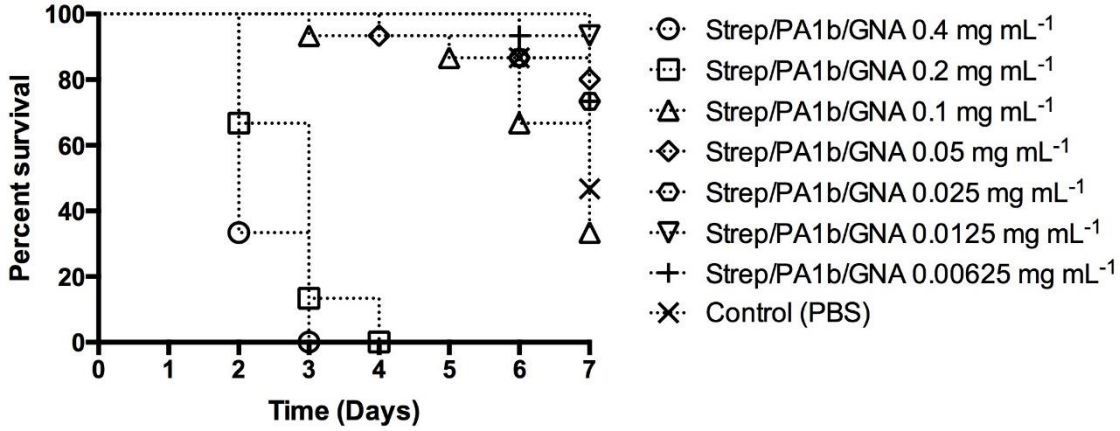


Figure 3.21 Survival of day one *A. pisum* nymphs fed on artificial diet containing recombinant Strep/PA1b/GNA. Purified recombinant proteins re-suspended in PBS were incorporated into artificial diet to give final concentrations as shown. The control treatment consisted of the diet with an equivalent volume of added PBS. N = 2.

3.7 Discussion

In this chapter the findings of Dr Pyati based around protein toxicity to aphids were independently confirmed, but a lack of activity of the GNA subunit and lack of purity prompted investigation into other fermentation methods, different constructs and a different purification method. Approaches taken in an attempt to enhance binding of Pa1b/GNA to nickel matrices included fermentation using different media in order to produce protein containing culture supernatants at a higher pH, new expressions constructs encoding an extended a linker region and an additional histidine residue were generated in an attempt to enhance binding of protein to nickel matrices. Unfortunately, none of these approaches were successful in recovering good quality protein. Finally, a small quantity of a relatively clean PA1b/GNA sample was eventually produced based on fermentation of a new expression construct that contained a strep-tactin affinity tag allowing purification by strep-tactin affinity chromatography. This sample was found to have substantially reduced efficacy as compared to samples produced by Dr Pyati suggesting that contaminants present in early samples were responsible for the high levels of toxicity that were initially recorded.

Previous work by Dr Prashant Pyati, revealed that fusion of PA1b and GNA may confer vastly increased orally active towards aphids, with doses as low as 0.05 mg mL^{-1} resulting in complete mortality in as few as four days (unpublished data). However, the purification process was arduous, involving dialysis of the culture supernatant then lyophilisation, resuspension in buffer followed by purification under denaturing conditions, dialysis and lyophilisation. The resulting lyophilised powders had low percentages of impure recombinant protein. The findings of toxicity toward *A. pisum* were replicated by feeding assay, and the correct sequence of the fusion protein product was confirmed by LC-MS analysis. The level of toxicity of PA1b/GNA toward *A. pisum* was substantially higher than that observed for other GNA fusion proteins, approximately 15-fold more toxic than PI1a/GNA and Pro-Hv1a/GNA, and much higher than reported in the literature for PA1b purified by reverse phase high purity liquid chromatography (HPLC) (Gressent *et al*., 2007; Yang, 2015; Yang *et al*., 2014). PA1b/GNA failed to agglutinate rabbit erythrocytes, suggesting that the GNA component of the fusion protein was non-functional. Previous work has shown that the transport of GNA to the haemolymph after feeding is reliant upon the ability of GNA to bind to the insect gut epithelium (Fitches *et al*., 2001). Therefore, if GNA is not active, enhanced toxicity of the PA1b/GNA fusion protein would not be expected, as anchoring of the protein to, or transport across, the epithelial membrane is unlikely to occur. It was thus hypothesised that perhaps confounding factors may be responsible for the observed toxicity of PA1b/GNA towards *A. pisum*. The data presented in this chapter describes

attempts to produce and purify a clean sample of a PA1b – GNA fusion protein. Once a relatively clean sample was produced the toxicity of that sample was assessed against *A. pisum*. A blank run control derived from performing the purification procedure on the supernatant of an untransformed yeast culture or a heat inactivated protein sample could have been used to confirm this hypothesis and should be tested the next time this particular issue and hypothesis is encountered.

Initially PA1b/GNA purification was attempted in the same manner as described above (Section 2.8.2). When samples of culture supernatant were purified by nickel affinity under native or denaturing conditions poor binding efficiency was observed although binding appeared to be slightly better under denaturing conditions. Unfortunately, purification under denaturing conditions followed by re-folding via dialysis using urea gradient did not result in the production of soluble protein in lyophilised powders. As insoluble protein was detected in pellets from centrifuged samples it is likely that the protein failed to refold during dialysis against a decreasing urea gradient. Perhaps on-column re-folding could be utilised if this were to be attempted again. Failure of the PA1b peptide to refold is not surprising, as literature suggests that PA1b may be dependent on the second protein encoded by the PA1 gene, PA1a, in order to fold correctly. However, due to its greatly enhanced toxicity compared to previous GNA fusion proteins and the lack of bioactivity displayed by PA1a the construct work continued with fusion constructs containing only PA1b and GNA to determine if the toxicity observed was a real effect or a result of confounding factors (Chouabe *et al*., 2011; Eyraud *et al*., 2013, Pyati - unpublished data) Constructs which contain PA1b and PA1a are examined in Chapter 5 and 6. Similarly, recombinant GNA when expressed in *E. coli* accumulates as insoluble inclusion bodies and re-folding by dialysis results in very low levels of recovery of folded, functional GNA (Fitches *et al*., 2001). It is likely that poor binding to nickel affinity columns was due the protonated state of the histidine tag in relatively low pH (approximately pH 5.0) loaded culture supernatant samples. It is known that optimal interaction between nickel and histidine occurs when the his-tag is deproteinated and carrying a negative charge and thus coordination between the Ni^{2+} and the histidine is much more probable (Bornhorst and Falke, 2000). Furthermore, with a pI of 5.78, it is likely, even when present at low concentrations in purified samples, that PA1b/GNA was prone to precipitation during dialysis against distilled water where the pH is raised from below the pI of the protein to approximately pH 6.5.

In an attempt to address issues pertaining to pI and pH, fermentations where cultures were grown at higher pHs were performed. Alternative media had to be used for higher pH fermentation as PTM1 salts, a key part of BSM, precipitate when fermentation media pH is raised

above 5.0. YPG and BMGY media are both utilised for fermentation of recombinant proteins in *P. pastoris* (Potgieter *et al*., 2009) (section 2.6) were used to culture PA1b/GNA at pH 6.5 and 6.0, respectively. In YPG medium low expression levels were observed and this low expression level led to low protein recovery and low purity of PA1b/GNA in lyophilised powders. A similar result was observed for fermentation in BMGY and whilst protein binding to the column was enhanced as compared to YPG, protein recovery was again poor, and this was likely to be due to precipitation during dialysis resulting in insoluble protein in the lyophilised powder. In addition to the protein of interest many extraneous were bands present in the final lyophilised powder. Purification by nickel affinity from medium containing eukaryotic proteins has been shown previously to result reduced purity of the final sample (Lichty *et al*., 2005).

The contamination observed in both fermentations is likely to do with the low protein content in the medium and poor column binding in combination with the yeast proteins present from the ingredients of the rich media. Normally the tagged protein should have a higher binding affinity for the nickel compared to non-tagged proteins, but given the low protein content and poor binding affinity of PA1b/GNA in the case of YPG fermented cultures together with the high binding capacity (40 mg mL^{-1}) of 5 mL Ni-NTA columns it is likely that binding sites were freely available for low affinity proteins (Bornhorst and Falke, 2000; Kimple *et al*., 2013). Bornhorst and Falke (2000) further suggested that non-specific purification may occur due disulphide bridge formation or hydrophobic interactions between the protein of interest and proteins present in the medium, which would facilitate co-purification. PA1b has both a high proportion of cysteine residues (6) and the active site of PA1b has an exposed hydrophobic surface, and as such both of the above explanations may be feasible (Da Silva *et al*., 2010; Eyraud *et al*., 2013). Purification with a smaller volume column or with additives such as low concentrations of Triton X-100 or β-mercaptoethanol may have proved beneficial as they reduce hydrophobic interaction and prevent disulphide bridge formation, respectively.

As purification under native conditions was a key priority new expression constructs were produced. The initial construct was the PA1b/GNA fusion protein but in the reverse orientation, GNA/PA1b which would result in the his-tag being in a different position relative to the orientation of PA1b and it was thought that this may increase tag availability and hence binding to nickel columns. Unfortunately, transformed *P. pastoris* clones failed to express GNA/PA1b. The complete lack of expression was surprising as GNA-toxin fusion proteins, such as GNA/Hv1a were expressed in *P. pastoris* (Powell *et al*., 2019).

As demonstrated previously issues pertaining to purification of PA1b/GNA were related to tag availability. Traditionally toxin – GNA fusion proteins have used a three alanine linker

region (Fitches *et al*., 2004; 2012; Powell *et al*., 2019; Yang *et al*., 2014). A new PA1b/GNA construct containing a flexible extended linker region was produced (gly-gly-gly-ser-ala-ala-ala) (PA1b/EL/GNA), based upon literature that suggests flexible linker religions allow correct folding of each subunit by increasing the space between the function proteins. It was also thought that an extended linker might serve to increase the distance between PA1b and the his-tag and prevent interaction. The GGGS linker has been shown to increase stability and folding along with biological activity of fusion proteins (Bai and Shen, 2006; Hu *et al*., 2004; Huston *et al*., 1988). PA1b/EL/GNA was expressed at comparable levels to PA1b/GNA when fermented in BSM. Recovery of this protein was poor (ca. 3.7 %) with contaminating proteins present in purified samples and this was again attributed to poor interaction between the his-tag and nickel columns.

As contamination and poor interaction with the columns were a continued issue a construct with an extended his-tag (7xH) was generated with the aim of improving binding efficiency. Previous work has shown that extension of the his-tag can improve the binding affinity of recombinant proteins without affecting function (Mohanty and Wiener, 2004). Expression of this construct was comparable to PA1b/GNA and PA1b/EL/GNA but unfortunately the extra histidine did little to improve binding to nickel columns and purity of the final sample was again poor.

Using the PA1b/EL/GNA construct as a template, a new construct incorporating an N-terminal strep-tag was generated. Evidence suggests that strep-tag purification often provides increased purity as compared to his-tag purification, particularly from yeast cultures (Lichty *et al*., 2005), and this was indeed the case. However, poor binding of the recombinant protein to strep-tactin columns was observed and may, as for his-tagged PA1b/GNA protein again have been due to limited accessibility of the Strep tag preventing binding to strep-tactin columns. Unfortunately, this could not be tested as strep-tactin affinity purification relies on a protein – protein interaction so denaturing purification was not possible. When run on SDS-PAGE the final product was significantly cleaner than previous attempts, with only one contaminating 45 kDa protein. This was in line with findings in the literature that report higher levels of purity resulting from strep-tactin affinity purified proteins as compared to nickel affinity purified proteins (Fang and Ewald, 2004; Lichty *et al*., 2005). LC-MS analysis of digested protein bands confirmed expression of intact Strep/PA1b/GNA whereas the 45 kDa band was found to be a complex mix of *P. pastoris* proteins. As this was the cleanest sample of a PA1b – GNA fusion protein produced Strep/PA1b/GNA was supplied in diet to *A. pisum* to assess its oral activity.

When nymphs of *A. pisum* were supplied with Strep/PA1b/GNA it was not as active as the original samples of PA1b/GNA. The LC_{50} was calculated to be 6.2 nmol mL^{-1} on day 5 as compared to 3.0 nmol mL^{-1} (estimated from data produced by Dr Pyati) originally obtained for PA1b/GNA. When considering molar concentrations of PA1b present in PA1b/GNA Despite reduced activity compared to initial samples of PA1b/GNA, Strep/PA1b/GNA was still more toxic to *A. pisum* than PA1b purified from seeds by HPLC. When 140 nmol mL^{-1} of HPLC purified PA1b was supplied in diet complete mortality occurred in six days as compared to the highest dose of Strep/PA1b/GNA tested, 23 nmol mL^{-1} which resulted in complete mortality after three days (Gressent *et al*., 2007). This evidence suggests that fusion of PA1b to GNA does increase toxicity compared to the PA1b alone.

The data in this chapter describes attempts to produce and purify – PA1b - GNA fusion protein under native conditions to provide pure and sufficient protein for further characterisation of insecticidal activity. This goal was unfortunately not achieved with this protein, despite the use of different fermentation media, the generation of new constructs containing extended linker regions and/or different purification tags. However, the production of sufficient Strep/PA1b/GNA protein of high purity did allow determination of efficacy against pea aphids, confirming that the very high levels of toxicity observed in preliminary assays were not attributable to the fusion protein itself and were more likely to be due to the presence of high concentrations of carbohydrate or contaminating proteins. An alternate composition of this fusion protein, which includes the second peptide sequence from the PA1 gene cassette will be investigated in the subsequent chapters.

Chapter 4 |Glucose or glycerol: alternative carbon sources for expression of *Galanthus nivalis* agglutinin (GNA) in *Pichia pastoris*.

4.1 Introduction

This study was prompted by issues encountered during the purification process of *Galanthus nivalis* agglutinin (GNA) and GNA based fusion proteins. As glycerol is fed continuously at an empirically determined rate that is increased manually throughout fermentation in bench-top vessels there is a possibility of residual glycerol being present in the final fermented culture medium. It was suspected that excess glycerol in the medium could lead to difficulties during purification of the recombinant protein; the protein is exported to the medium post expression due to the α-factor secretory signal in the expression vector. Specifically, glycerol may be interfering with the interaction between the nickel and the 6xhistidine residue tag, reducing the binding affinity of the recombinant protein. However, peer-reviewed literature pertaining to the effect of glycerol on purification by nickel column chromatography is lacking.

Pichia pastoris can grow on many carbon substrates. Glycerol, glucose and methanol are the most popular and widely used (Inan and Meagher, 2001). In this experiment, the glyceraldehyde-3-phosphate dehydrogenase (GAP) promoter was used to drive expression; therefore, methanol was discounted as it was not necessary for protein expression, and it is toxic and flammable. Glucose was selected in this instance as an alternative carbon source as it may have beneficial effects on recombinant protein expression level and does not affect the specific growth rate of *P. pastoris* fermentation (Waterham *et al*., 1997; Rebnegger *et al*., 2014; Prielhofer *et al*., 2015). Furthermore, it has a comparable metabolic pathway to glycerol unlike methanol; glycerol is catalysed by a number of enzymes to glyceraldehyde-3-phosphate, from which point it joins the glycolysis pathway and continues as glucose would (Vanz *et al*., 2012). GNA was selected as a candidate protein for this experiment as it is well characterised and has been produced in *P. pastoris* for many years, and is one of the proteins where problems with poor recovery from culture supernatant, due to inefficient binding to the nickel affinity columns were encountered (Baumgartner *et al*., 2003; Raemaekers *et al*., 1999).

4.2 Expression construct for recombinant GNA

The *P. pastoris* clone used to express GNA was produced by Dr Prashant Pyati. The sequence was codon optimised for *P. pastoris* expression and synthesised commercially (Shinegene). The construct was subsequently ligated into the pGAPzαB expression vector via *Pst*I and *Sal*I restriction sites. The DNA and derived peptide sequences from a sequenced clone are shown in Figure 4.1.

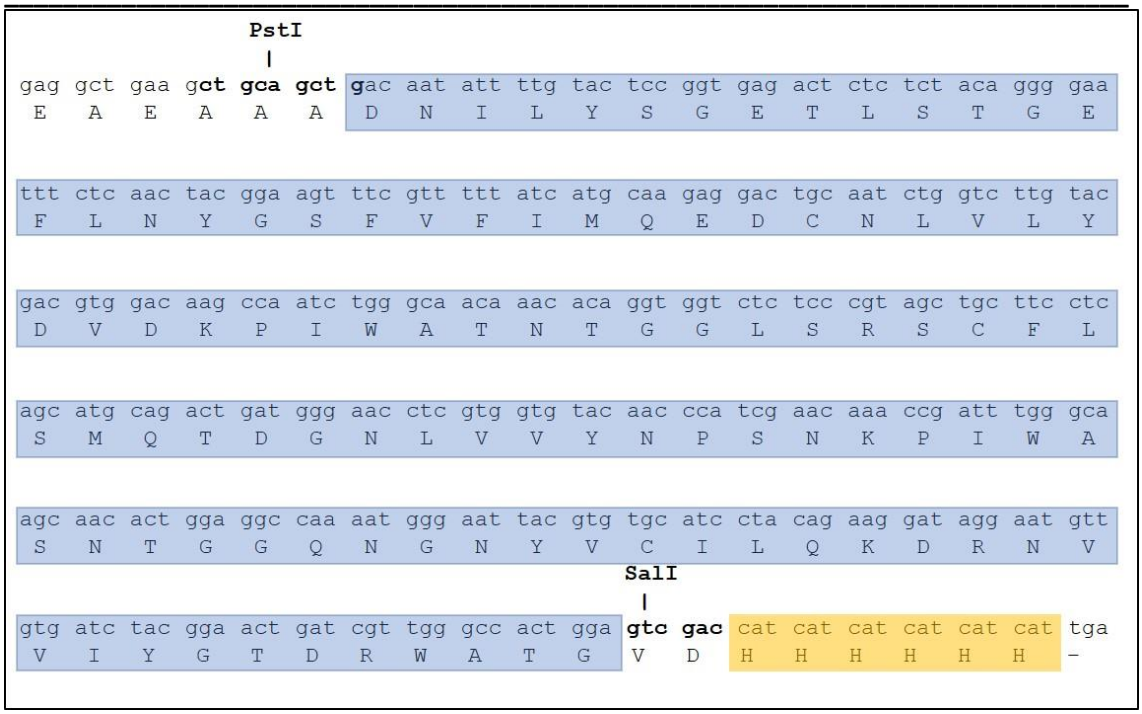


Figure 4.1 GNA nucleotide and derived peptide sequence from a sequenced expression construct. Generated by Sanger sequencing of a positive clone. GNA peptide sequence shown in blue and the 6xHis peptide sequence shown in yellow. Restriction sites are shown in bold. Predicted MW: 13.2 kDa. Predicted pI: 5.08.

4.3 Expression and purification of GNA fermented with different carbon sources

The GNA expressing clone was inoculated for both fermentations into 3 x 100 mL flasks containing yeast peptone glycerol medium (YPG) and incubated at 37 °C for 72 hours. The two flasks with the best growth (assessed visually) were selected and inoculated into a 7 L benchtop bioreactor with basal salt medium (BSM). The protein was fermented under controlled conditions listed in section 2.6. The fed carbon source was modified, with 50 % (v/v) glycerol used in one instance and 50 % (w/v) D-glucose used in the other. The initial carbon source in the BSM was also adapted to which ever carbon source was fed in that fermentation.

Wet pellet weight was monitored throughout both fermentations as a proxy for cell growth (Table 4.1). Wet pellet weight increased throughout both fermentations. At time 0 wet pellet weights were 20.4 mg mL^{-1} and 15.3 mg mL^{-1} for glycerol and glucose, respectively, increasing to 321.7 mg mL^{-1} and 217.8 mg mL^{-1} after 72 hours of fermentation.

When the cell pellet was collected the supernatant was sampled and expression of protein was assessed in both fermentations by western blotting with anti-GNA antibodies. As shown in Figure 4.2, an immunoreactive band was visible at approximately 14 kDa in fermentation supernatants. This was 1 kDa greater than the predicated mass of GNA, likely as a result of glycosylation (Macauley-Patrick *et al*., 2005). Band intensity increased throughout both fermentations despite decreased loading volumes (to prevent overloading). Final expression level was estimated semi-quantitively by western blot, comparing the intensity of immunoreactive bands in the supernatant to standards of known quantity. The expression of GNA was approximately 100 mg L^{-1} when cultures were fed glycerol, and approximately 40 mg L^{-1} when fed glucose (Figure 4.2).

After 72 hours the culture was harvested, and the supernatant was separated from the cells by centrifugation and filtered to remove any remaining cells. The recombinant GNA was purified from glycerol and glucose fed culture supernatants by nickel affinity chromatography under identical conditions, following the protocol described in section 2.8. In this instance 450 mL of each supernatant was loaded onto 2 x 5 mL HisTrap FF columns for 2 hours. All fractions were collected and analysed by western blot. In glycerol-fed samples (Figure 4.3**Error! Reference source not found.**A), the intensity of the immunoreactive band was similar in both the supernatant and the flow through, suggesting poor binding affinity to the nickel column. By contrast, GNA fermented in glucose-fed medium bound to the columns much more efficiently (Figure 4.3B); the flow through band intensity was substantially reduced compared to the supernatant. This indicates a potential detrimental effect of glycerol on nickel-histidine

interaction. GNA was detected in the eluted peak fraction of both fermentations but not the 10 mM imidazole wash steps.

Table 4.1 Wet pellet weight, from the production of recombinant GNA using glycerol or glucose as a carbon source. Sample was removed from the bioreactor every 24 hours after inoculation and then at the end of the fermentation process.

	Glycerol	Glucose
Time (h)	**Wet pellet weight (mg mL^{-1})**	**Wet pellet weight (mg mL^{-1})**
Inoculum	82.1	50.8
0	20.4	15.3
24	119.4	69.7
48	314.5	216.7
72	321.7	217.8

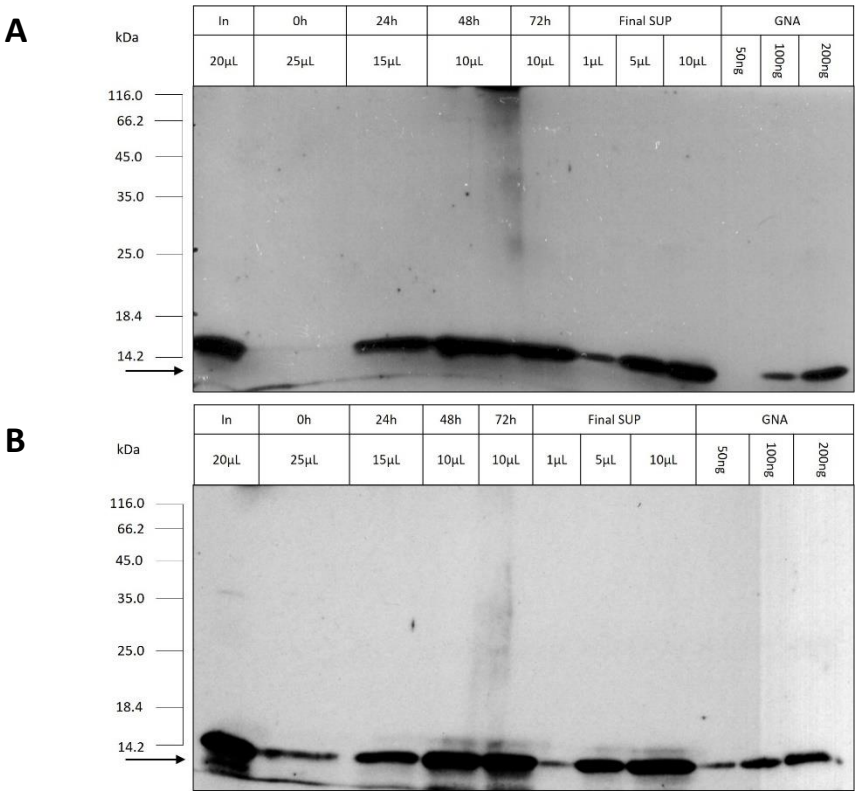


Figure 4.2 Western blot (anti-GNA antibodies), of culture supernatants from recombinant GNA *Pichia pastoris* fermented with glycerol (A) and glucose (B). Culture was removed from the bioreactor after inoculation and then every 24 hours until the end of the fermentation (72 hours). Time points and volume loaded in microlitres are stated above the lanes. In: inoculum, SUP: Supernatant, GNA: recombinant GNA standards, quantity stated above the line in nanograms A: 50 % (v/v) glycerol used as a carbon source. B: 50 % (w/v) glucose used as a carbon source. The arrow indicates the expected protein mass.

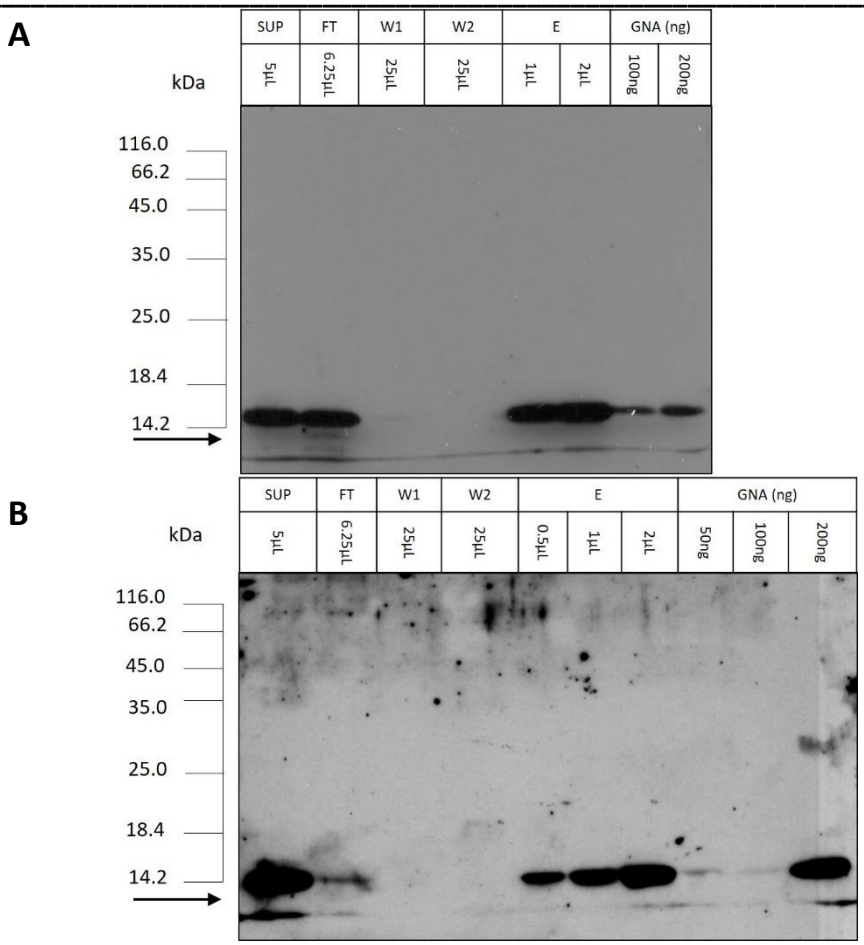


Figure 4.3 Comparison of binding efficiency of GNA to Ni-NTA resin by western blot analysis using Anti-GNA antibodies. Loaded fraction and volume in microlitres are stated above the lane. SUP: supernatant, FT: Flow through, W1: 1xBB, W2: 1xBB + 10 mM imidazole, E: 1xBB + 200 mM imidazole, GNA: recombinant GNA , quantity stated in nanograms. A: GNA from glycerol fed fermentation, B: GNA from glucose fed fermentation. The arrow indicates the expected protein mass.

To further test the effect of glycerol on binding efficiency, increasing volumes of glycerol were added to 100 mL of supernatant derived from glucose fed medium to determine directly whether the addition of glycerol decreases the binding of recombinant GNA to the nickel column. Shown in Figure 4.4, the load (supernatant + glycerol + binding buffer) and flowthough was analysed at each glycerol concentration by western blot in order to assess binding efficiency under differing concentrations of glycerol. At low glycerol concentrations, 0 % and 0.75 % (v/v), there was a decrease in band intensity between the load and flowthough. At high concentrations there was no decrease in band intensity between load and flow through. This finding was confirmed in the analysed elution fractions from each purification (Figure 4.4).

After purification the eluted fractions were de-salted by dialysis and then lyophilised as described in section 2.10. To determine the proportion of GNA in the lyophilised powder known quantities of powder were compared to known quantities of commercial GNA standards by total protein staining of an Sodium dodecyl sulphate polyacrylamide gel elcotrophoresis (SDS-PAGE) gel, shown in Figure 4.5. GNA derived from glycerol-fed medium made up approximately 20 % (w/w) of the lyophilised powder, whilst GNA derived from glucose-fed medium made up 15 % (w/w). The total protein recovery from each fermentation was 39 % and 51 % respectively. The lower recovery rate of GNA provides further evidence for the hypothesis that glycerol may be reducing the binding efficiency of the nickel column.

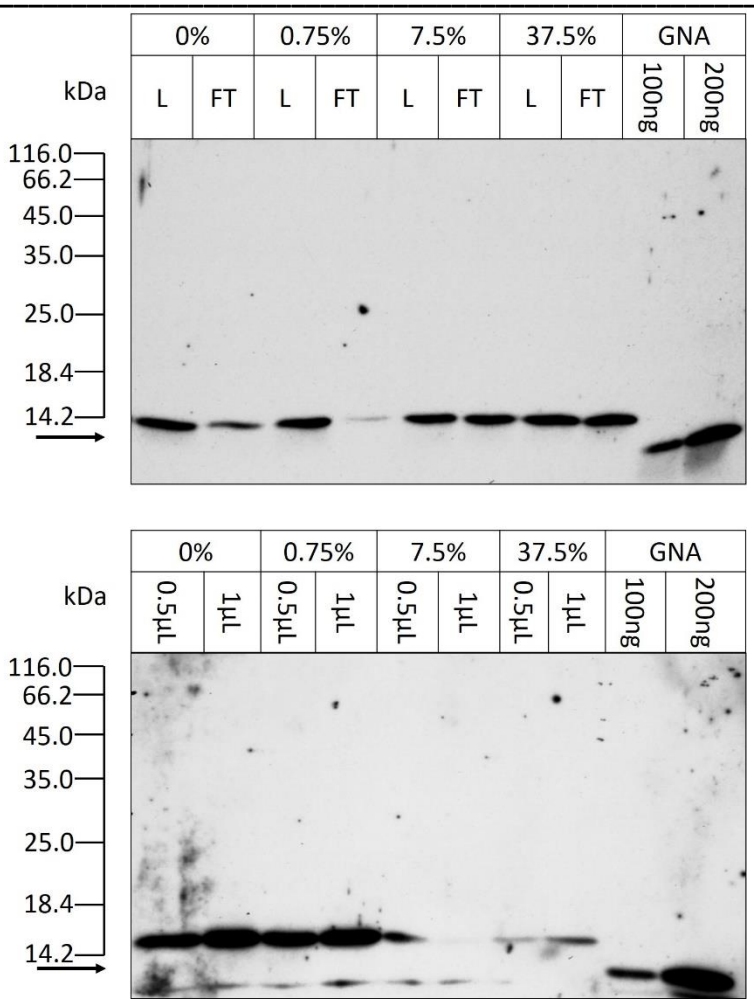


Figure 4.4 Comparison of Load and flow through (top) and elution fractions (bottom) with artificially added glycerol. Western blot probed with Anti-GNA antibodies. Each sample was passed through a Ni-NTA affinity column for 1 hour. Percentages represent the concentration of glycerol (v/v) in the final loaded sample. Recombinant GNA standards A: L: Load, FT: Flow through, Equal volumes were loaded in all lanes. B: Comparison of 200 mM Imidazole + 1 BB eluted peak fraction, volume (µl) stated above the lanes. The arrow indicates the expected protein mass.

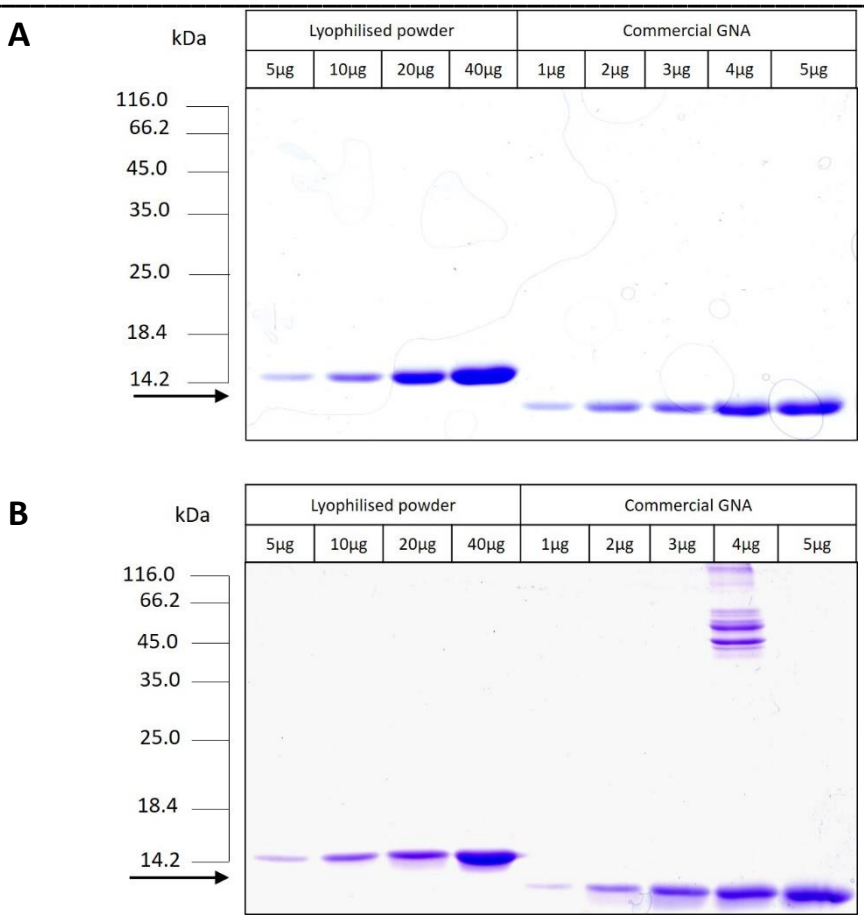


Figure 4.5 Quantification of GNA proportion in lyophilised powder derived from glycerol-medium (A) and glucose-fed medium (B). Coomassie brilliant blue stained SDS-PAGE gel (17.5% acrylamide). Quantity of lyophilised powder and commercial GNA standards (Sigma-Aldrich, St. Louis, USA) are stated above the lanes. The arrow indicates the expected protein mass.

4.4 Biological analysis of GNA fermented with either glycerol or glucose as a carbon source.

Glycosylation of the recombinant GNA samples was assessed by Schiff-periodic acid staining as described in section 2.13. This was done as *P. pastoris* has been shown to glycosylate recombinant proteins during expression, and this can affect the molecular mass and on occasion function (Macauley-Patrick *et al*., 2005). There are two types of glycosylation, N-linked and O-linked. N-linked glycosylation requires a specific N-X-S/T amino acid motif (Schwarz and Aebi, 2011). There are no N-linked motifs present in the amino acid sequence of the construct used (Figure 4.1). A glycosylated band was detected at approximately 14 kDa, where expected. No difference was visible between the two GNA samples stained with Schiff reagent (Figure 4.6). Recombinant PI1a/GNA, which contains a single N-X-S/T amino acid motif, was included as positive control, and commercial GNA (purified from *G. nivalis* bulbs) was used as a negative control (Van Damme *et al*., 1987; Yang *et al*., 2014).

The activity of GNA, and thus the folding of GNA, was assessed by a hemagglutination assay with rabbit erythrocytes. GNA produced using either glucose or glycerol as a carbon source was able to agglutinate rabbit erythrocytes down to a threshold concentration of 15.6 µL mL^{-1} (Figure 4.7). This was comparable to recombinant GNA produced by Dr Min Cao independently of this experiment (unpublished data). GNA purified from *G. nivalis* bulbs (Sigma-Aldrich, St. Louis, USA) used as a positive control was found to agglutinate down to a threshold concentration of 7.8 µL mL^{-1}. PBS used as a negative control, showed no agglutination activity. The threshold for agglutination was determined visually.

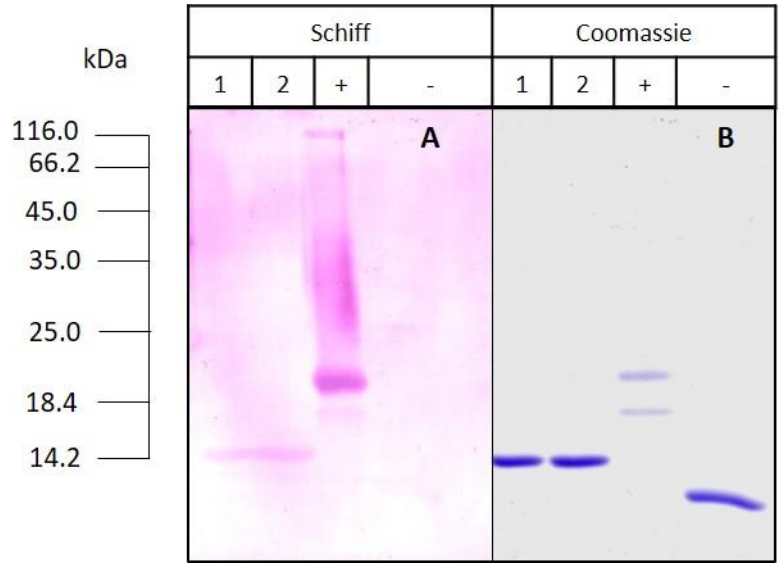


Figure 4.6 Analysis of glycosylation status of recombinant GNA fermented in glycerol fed medium or glucose fed medium. A: Schiff reagent stained nitrocellulose membrane, 10 µg of protein loaded in all lanes. B: Coomassie brilliant blue stained SDS-PAGE gel (17.5 % acrylamide) 5 µg of protein loaded. 1: GNA (Glycerol), 2: GNA (Glucose), +: PI1A/GNA, -: GNA (Purified from snowdrop, Sigma-Aldrich, St. Louis, USA)

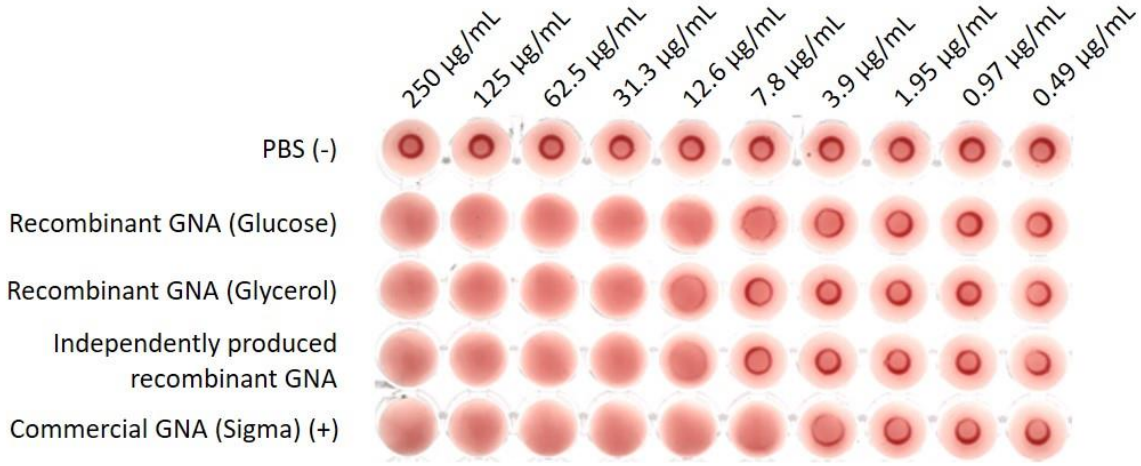


Figure 4.7 Agglutination assay to determine the mannose binding activity of GNA. Protein was serially diluted and incubated with a 2 % rabbit erythrocyte solution. After 2 hours the plate was imaged to determine the lowest concentration at which GNA can crosslink rabbit erythrocytes. Agglutinated or crosslinked cells display an even red colour in the well whilst non-agglutinated cells settle to the bottom of the plate and form a visible dark red pellet.

4.5 Discussion

Evidence is provided in this chapter that demonstrates when glycerol is present in the supernatant in high concentrations, there is a detrimental effect on purification of recombinant GNA. Changing the carbon source to glucose has no detrimental effect on growth, purification or biological activity of the recombinant GNA, and has enhanced recovery via more efficient binding to the nickel columns as compared to glycerol containing supernatants.

Pichia pastoris is a methanotrophic yeast which grows under several different carbon regimes (Waterham *et al*., 1997). In this chapter both glycerol and glucose were used as carbon sources to determine if recombinant GNA recovery could be improved in fermentation medium lacking glycerol. Growth rate was tracked, through the monitoring of wet pellet weights in both fermentations, to determine if the carbon source affects growth rate and therefore protein expression (Rebnegger *et al*., 2014). The difference in final wet pellet weight observed is unlikely to be due to the difference in carbon source as existing literature suggests that the growth rate for *P. pastoris* is unaffected by the use of glycerol or glucose (Prielhofer *et al*., 2015). To corroborate the literature the specific growth rate of both conditions was calculated using the wet pellet weights from time 0 and 72 hours to give the increase in mg of wet pellet weight per mL per hour using the equation described in (Rodrigues *et al*., 2019). *Pichia pastoris* grown in glucose-fed medium had a specific growth rate of 0.037 h^{-1}. The specific growth rate in glycerol-fed medium was slightly higher (0.038 h^{-1}), indicating that *P. pastoris* was growing at a comparable rate in both conditions. Prielhofer *et al*., (2015) quotes a specific growth rate for fed-batch cultures between 0.010 and 0.022 h^{-1}, which is comparable, although lower than values found in this experiment. The discrepancy between the two results could be due to the medium used in the culture (YPG versus BSM) and the culture method (shake flasks with slow release glucose disks versus bioreactor fermentation). It is important to note that due to the fed-batch nature of benchtop fermentation, where additional dissolved carbon is added to the vessel, the specific growth rate is not entirely representative of growth in the culture as it assumes there is no change in culture volume. However, as equal volumes were added the specific growth rates of the two fermentations can be compared. This difference in final pellet weight is likely due to a reduced inoculum pellet weight. The inoculum glycerol wet pellet weight was 82.1 mg mL^{-1}, whilst the inoculum glucose wet pellet weight was 50.8 mg mL^{-1}.

Under both carbon regimes the protein expression profile increased over time, in much the same manner as the increase in wet pellet weight (Table 4.1,Figure 4.2). Comparison of supernatant to known quantities of protein standard allowed visual estimation of the final expression level in the medium. The glycerol-fed supernatant had an estimated expression level

of approximately 100 mg L^{-1} whilst the glucose-fed supernatant expression level was estimated at 40 mg L^{-1}. This marked decrease in expression level is contradictory to published literature which suggests that recombinant proteins under the control of the GAP promoter should be expressed at a higher rate when glucose is provided in the medium compared to glycerol (Waterham *et al*., 1997). This difference in protein expression level between the two fermentations again may be due to the decreased inoculum cell density, as heterologous protein expression is linked to both specific growth rate and cell density (Rebnegger *et al*., 2014). As specific growth rate was found to be similar in both fermentations, the difference in expression level may be due to decreased inoculum cell density, resulting in a lower final cell density. The final wet pellet weight of the glucose fermentation was 66 % of the glycerol fermentation wet pellet weight, while the expression level was 40 %, and as the expression levels are approximations these values can be considered comparable. It is important to consider that this experiment was only performed once, the conditions for the glucose fermentation have not been optimised, and expression level can vary from fermentation to fermentation.

Nickel affinity chromatography functions by forming a coordinate covalent bond between an immobilized Ni^{2+} ion and two histidine side chains. In this case the protein has been expressed with a N-terminal 6x polyhistidine tag to allow purification (Figure 4.1). From discussion with industrial collaborators, and anecdotal evidence from experience with the *P. pastoris* protein expression system, it was hypothesised that excess glycerol in the medium was interfering with the interaction between the nickel ion and on the polyhistidine tag. There is little to no published literature on this topic, but glycerol is listed by the manufacturer (HisTrap FF, GE Healthcare) as a compatible compound in concentrations up to 50 % (w/v).

The effect of glycerol on purification by nickel affinity chromatography was tested by a comparison between two fermentations, one containing glycerol as a carbon source and one containing glucose as a carbon source. When glucose was used as a carbon source less recombinant protein was left in equal volumes of supernatant after two hours loading on nickel columns compared to the supernatant from the glycerol-fed fermentation (Figure 4.3). These results in conjunction with reduced binding and recovery of GNA from nickel affinity columns observed when increasing levels of glycerol were added to loaded supernatants (derived from a glucose fed fermentation run) provides evidence that concentrations of greater than 7.5 % (v/v) glycerol have a negative impact upon GNA yield. There did appear to be a beneficial effect of 0.75 % (v/v) glycerol on GNA – nickel binding, perhaps suggesting a threshold concentration above which glycerol becomes detrimental during the purification of GNA. It is not surprising that low concentrations of glycerol could have a positive effect on purification as some

purification protocols call for low concentrations of glycerol (Bornhorst and Falke, 2000). The data from this experiment contradicts the compatibility list from the manufacturer, but there are many factors that have a profound effect on protein purification, and optimal conditions for purification must be determined experimentally on a case by case basis.

Based on the evidence gathered, glycerol appears to be detrimental to GNA protein binding above a certain threshold, however the exact mechanism is not clear. The column manufacturers (GE Healthcare) state that glycerol can reduce binding of nonspecific hydrophobic protein to nickel columns. This is possibly due to the tendency of glycerol to aggregate around hydrophobic regions of the proteins, this is unlikely to be the reason for the decrease in binding as GNA is a hydrophilic protein, and histidine is a polar amino acid. It is possible that this is purely a steric affect, based on the increased viscosity of the supernatant containing high concentrations of glycerol. A speculative explanation may be that although glycerol is not molecularly similar to the histidine side chain or imidazole, it is forming complexes with nickel. Nickel can form complexes with many compounds including the negatively charged hydroxyl groups of which glycerol has 3 located in proximity, whereas the 5 hydroxyl groups in glucose are dispersed around the 6-carbon ring. As the isoelectric point of GNA (5.08) was above the pH of the supernatants from which GNA was purified (approximately 4.9), the histidine residues were likely in a protonated state and less likely to interact with the positively charged nickel, allowing the glycerol to compete to form complexes. Furthermore, glycerol is half the mass of glucose, 92 Da compared to 180 Da and this may allow more glycerol to cluster around individual nickel ions. Therefore, it may be possible that glycerol is complexing with nickel and preventing the histidine from binding when present in high concentrations. It has also been shown that solutions of glycerol and water reduce nickel (II) complex formation between nickel and pyridine 2-azo-p-dimethylaniline at room temperature as compared to pure water, a similar effect may be occurring here between the nickel and the his-tag, although the histidine side chain is not similar to the molecule in this study (Ellis *et al*., 1989). Ultimately, the mechanism is not clear, in this case, the use of glucose as compared to glycerol as a carbon source for *P. pastoris* growth resulted in a 29 % increase in relative protein quantity being recovered from fermented cultures.

The use of different carbon sources had no discernible effects on the degree of glycosylation of recombinant GNA or its ability to cross link rabbit erythrocytes suggesting comparable *in vitro* functionality. Schiff staining confirmed glycosylation of recombinant GNA and, as there are no N-linked glycosylation sites present in the protein sequence of the

recombinant GNA it is highly likely that the glycosylation visible in the Schiff assay is O-linked (Figure 4.6).

Recombinant GNA, from both fermentations, agglutinated at a threshold of 15.6 µg mL^{-1}; this is approximately two times higher than published literature, whereas the positive control (GNA purified from snowdrop bulbs) was consistent with the positive control values in the same study (Raemaekers *et al*., 1999). This inconsistency could be attributed to several factors. This study used a truncated GNA peptide of 105 residues, versus the native 109 residue protein used by Raemaekers et a. (1999). The difference in length may interfere with the ability of recombinant GNA to form tetramers through C-terminal strand exchange which involves residues 99-109 (Hester *et al*., 1995). Tetramerization could account for the difference in agglutination between these experiments as it increases the number of available mannose binding sites per molecule from 3 to 12 allowing GNA to interact with more mannose residues per erythrocyte, enabling a stronger cross linking reaction (Hester *et al*., 1995). This is possibly why GNA purified from snowdrop could agglutinate down to a lower threshold than recombinant GNA. Other factors that could have a contributing affect could be the age of the erythrocytes, the amount of time the plates were left to settle and the carbohydrate content of the recombinant protein.

It is important to note that the results are preliminary as they are derived from single fermentation runs to compare the influence of different carbon sources upon the protein yield and no statistical evaluation can be performed. Furthermore, no optimisation had taken place related to carbon feed rates using glucose and it was assumed that glucose and glycerol fed at the same rate would provide similar energetic content. This was done to determine if the recovery of GNA could be improved using a different carbon source. If this work were to be repeated, ensuring inoculum cell densities were equal would be important to consider, along with optimisation of the carbon source feed rate.

In conclusion, there appears to be no difference between glucose and glycerol when used as a carbon source in the fermentation of a GNA expressing *P. pastoris* clone. There is some evidence to suggest that excess glycerol in the culture supernatant has a detrimental effect on the binding efficiency of nickel affinity columns and their ability to interact with poly-histidine tagged proteins. An interesting line of investigation would be to determine whether other proteins, especially those with poor binding affinity to nickel, exhibit similar behaviour when the carbon source is swapped to glucose instead of glycerol.

Chapter 5 |Production and purification of recombinant pea albumin (PAF) and a pea albumin based fusion protein (PAF/GNA) in *Pichia pastoris*

5.1 Introduction

Pea albumin, PA1b, is a small proteinaceous inhibitor of V-ATPase proton pumps, initially purified from the seeds of *Pisum sativum* (Muench *et al*., 2014). It was hypothesised that fusion of this small 36 amino acid protein to *Galanthus nivalis* agglutinin (GNA) could enhance its toxicity to aphids; using the binding affinity of GNA for mannosylated gut epithelial proteins to anchor PA1b in proximity to its target site. Unfortunately, many issues were encountered with the production and purification of PA1b/GNA (discussed in Chapter 4). Therefore, this chapter focuses on the production of recombinant full length pea albumin (PAF) and PAF/GNA. The PAF expression construct contains the coding sequences for both the PA1b polypeptide and the PA1a polypeptide from the *Pisum sativum* gene cassette PA1 (Higgins *et al*., 1986). These constructs were developed because evidence suggests that whilst PA1a is not toxic to Sf9 cells (Insect cell line from *Spodoptera frugiperda*) it is important for the functional expression of PA1b (Chouabe *et al*., 2011; Eyraud *et al*., 2013).

Fermentation conditions must be optimised on a case-by-case basis for *P. pastoris* protein expression. Therefore, this chapter presents data and discusses the production and optimisation of recombinant PAF and PAF/GNA under native purification conditions in basal salt medium, specifically looking into the effects of two different pHs and two carbon sources on expression levels and purification profiles of the final proteins. Glycosylation and agglutination activity, where appropriate, were also examined to determine if changes to the fermentation protocols affect activity of the GNA carrier or significantly change the quantity of glycosylation.

5.2 Expression constructs

The expression construct for recombinant pea albumin (PAF) was designed to be as similar as possible to the native protein (Higgins *et al*., 1986). The fusion protein consists of PAF linked to the n-terminus of GNA. The PA1 gene construct and mature GNA sequences were codon optimised for expression in *P. pastoris* (https://eu.idtdna.com/CodonOpt) and synthesised commercially. Restriction sites were added as necessary using PCR, and restricted products were subsequently ligated into pGAPzαB in frame with the α-factor secretory signal and the C-terminal histidine tag (his-tag), which allows detection by western blotting and purification by nickel affinity chromatography. Ligated plasmids were then transformed into TOP10 *E. coli* following the protocol described in section 2.2.9 and verified by colony PCR and DNA sequencing, then confirmed using Serial Cloner 2.6. The nucleotide and derived peptide sequences for PAF and PAF/GNA are shown in Figure 5.1 and Figure 5.2.

5.3 Screening for protein expressing clones

Vectors containing the sequenced expression cassettes were linearized with AvrII and transformed into *P. pastoris* SMD1168H as described in Sections 2.2.13 and 2.3. *Pichia pastoris* SMD1168H colonies which grew on zeocin selective medium, were inoculated into flasks containing 10 mL yeast peptone glycerol medium (YPG) and 100 µg mL^{-1} zeocin. The inoculated flasks were incubated at 30 °C for 48-72h. Supernatant samples were subsequently screened by western blot to select positively expressing clones. In Figure 5.3A, PAF expressing *P. pastoris* clone 4 showed a single immunoreactive band at approximately 10 kDa, slightly lower than the expected mass of 11.6 kDa. This clone was selected for over expression in a benchtop bioreactor as described. In small scale cultures of PAF/GNA transformants, an immunoreactive band was visible at approximately 25 kDa, 1 kDa higher than expected for PAF/GNA. A faint GNA immunoreactive band was also detected at 14 kDa, likely to be cleaved GNA. Clone 1 was selected as the best PAF/GNA expressing clone (Figure 5.3B).

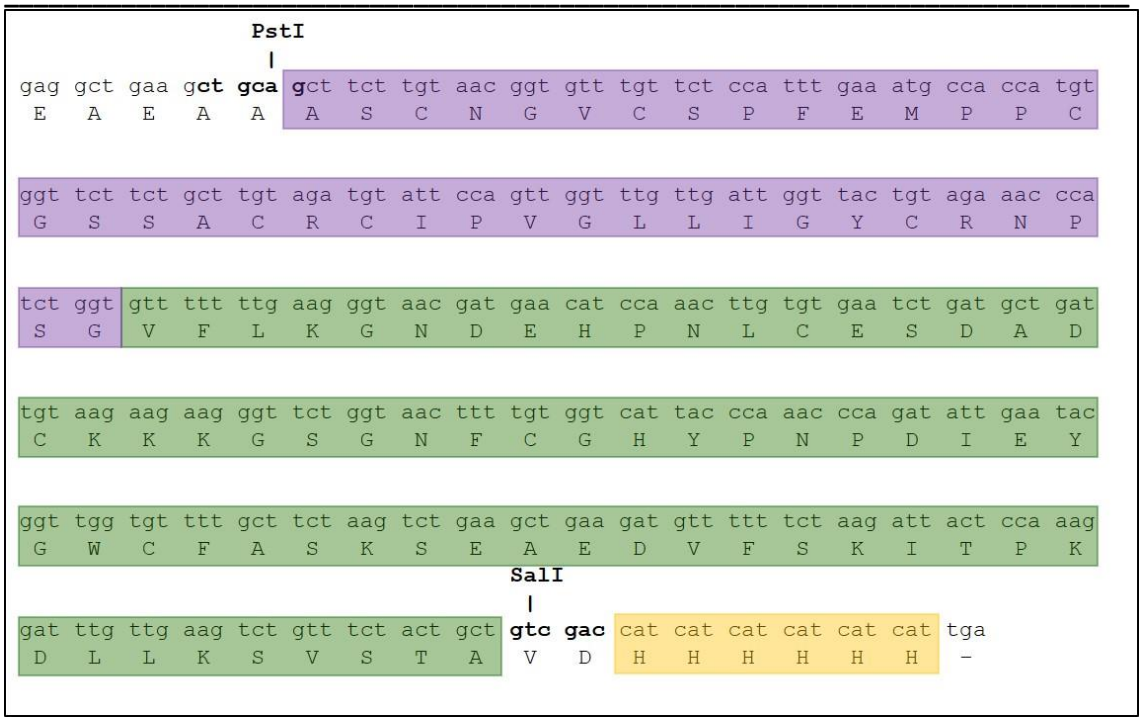


Figure 5.1 PAF nucleotide and derived peptide sequence, from a sequenced clone. PA1b (purple), PA1a (green), his-tag (yellow). Bold: Restriction sites used in cloning. Predicted Mw of peptide: 11.8 kDa.

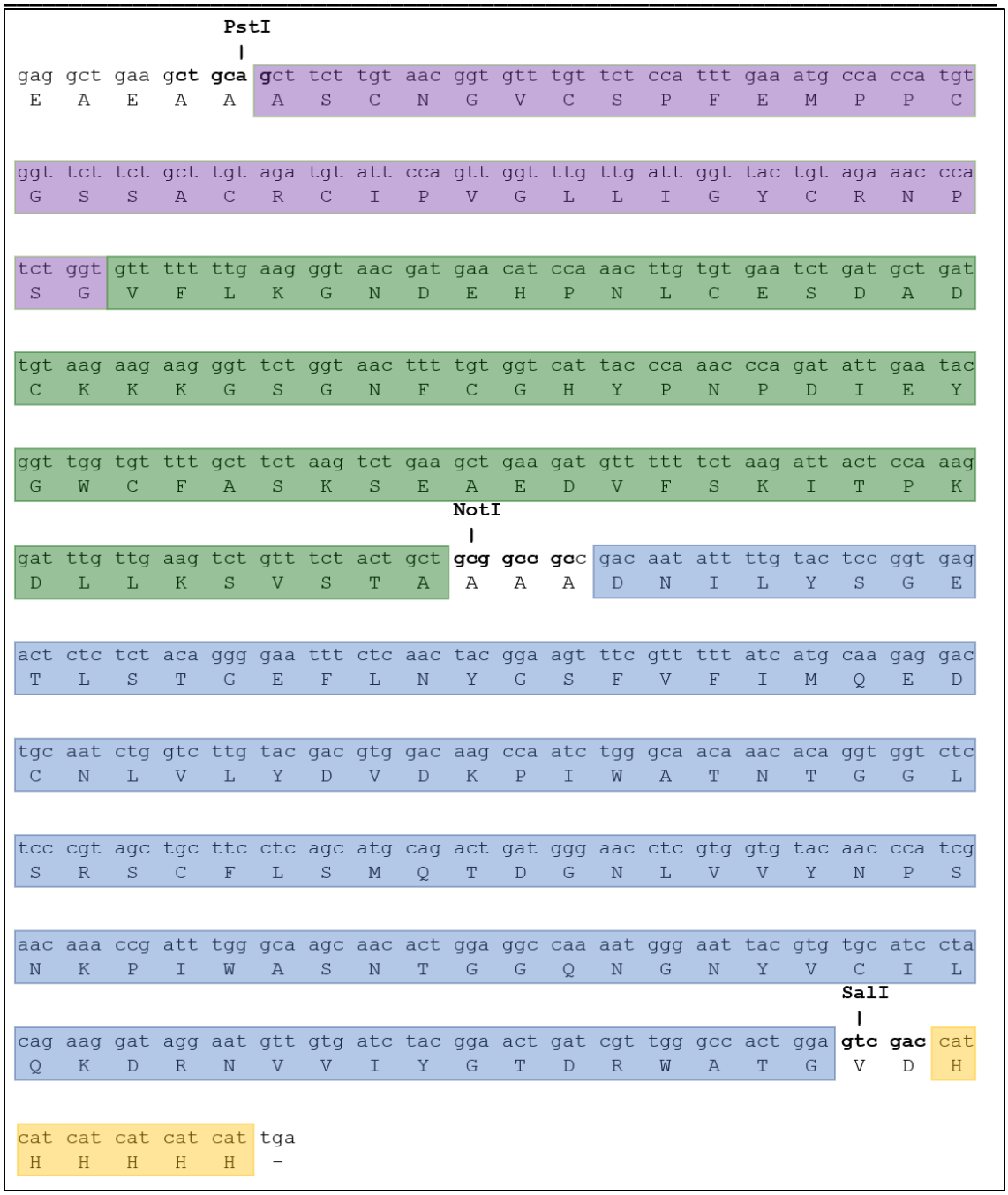


Figure 5.2 PAF/GNA nucleotide and derived peptide sequence, from a sequenced clone. PA1b (purple), PA1a (green), GNA (blue), his-tag (yellow). Bold: Restriction sites used in cloning. Predicted Mw of polypeptide: 23.9 kDa.

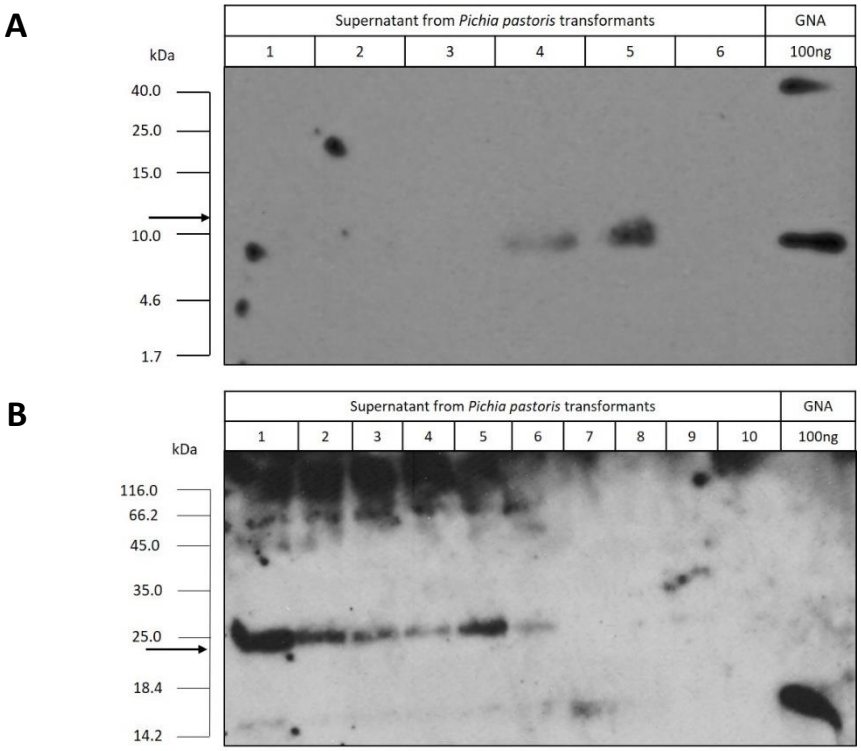


Figure 5.3 Screening of small scale *P. pastoris* cultures for the expression of recombinant PAF (A) and PAF/GNA (B). Western blot, probed with anti-his (A) and anti-GNA (B), screening to select PAF and PAF/GNA expressing *P. pastoris* SMD1168H clones. Individual colonies were inoculated into 10 mL YPG flasks with 100 µg mL^{-1} zeocin, incubated for 72h at 30 °C with shaking. Cultures were centrifuged and supernatant (25 µL) was loaded in each lane. Numbers denote individual colonies. GNA denotes 100 ng recombinant GNA that contains a C-terminal histidine tag. S1-S6 denotes supernatant from individual cultures. Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

5.4 Expression and purification of recombinant pea albumin (PAF) and pea albumin fusion protein (PAF/GNA)

Recombinant PAF was produced in a 7 L benchtop bioreactor, as described in section 2.6, using basal salt medium (BSM); the carbon source used was glycerol and pH was maintained by addition of 17.5 % (v/v) ammonia. Changes made to the protocol in section 2.6 are as follows: in one instance pH was maintained at 4.5, in the other pH was maintained at 5.0. The inoculum used to seed the bioreactor was grown in 3 x 100 mL YPG flasks at 30 °C for 72 hours.

PAF/GNA was produced in the same bioreactor, with the following changes made to that protocol in separate fermentations: A) No change: pH was maintained at 4.5 and the carbon source utilised was glycerol; B) pH was maintained at 5.0 and the carbon source used was glycerol; C) The pH was maintained at 4.5 and the carbon source was changed to glucose. The inoculum used to seed the bioreactor was grown for 72 hours at 30 °C in 3 x 100 mL YPG flasks.

5.5 Wet pellet weight of PAF expressing *P. pastoris* fermented at pH 4.5 and 5

The wet pellet weight was monitored throughout both PAF benchtop fermentations. Samples were taken from the fermenter at 24h intervals to monitor growth of the culture and expression of the recombinant protein. A comparison of wet pellet densities recorded for fermentation of PAF is shown in Table 5.1. The pellet weight at time 0, immediately post inoculation, showed little difference between pH 4.5 and 5 (16.4 mg mL^{-1} and 18.8 mg mL^{-1} respectively). The final wet pellet weights for these two fermentations were 219.8 mg mL^{-1} for pH 4.5 and 278.9 mg mL^{-1} for pH 5.0. Wet pellet weights were not collected for PAF/GNA fermentations.

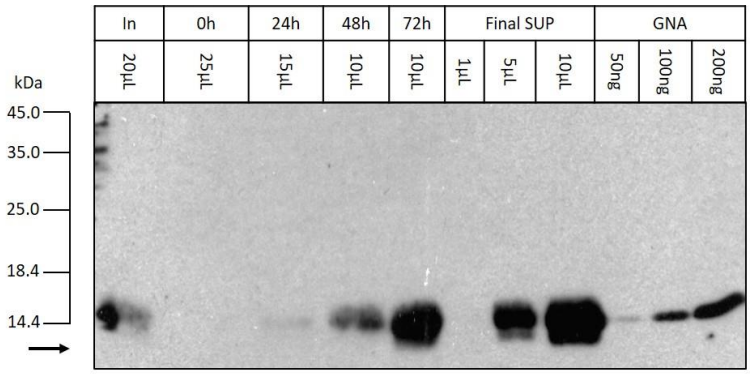
5.6 Estimation of expression level of PAF and PAF/GNA fermented at different pHs and using different carbon sources

Protein expression levels during, and at the end of fermentation, were assessed by western blotting of culture supernatants using anti-His (PAF) and anti-GNA (PAF/GNA) antibodies. Final expression levels in harvested cultures were estimated by visual comparison of the immunointensity of recombinant GNA standards with three different volumes of the final supernatant as shown in Figure 5.4A and B. An immunoreactive band of approximately 14 kDa is present in samples from fermentations conducted at pH 4.5 and pH 5.0 (Figure 5.4). The 14 kDa band is 2.2 kDa greater than the predicted molecular weight of PAF, this increased mass is likely to be due to glycosylation. Band intensity increased over time, suggesting the quantity of recombinant protein increased as the cultures grow. The results presented in Figure 5.4 correlate with the wet pellet weights (Table 5.1). The final expression level of PAF was estimated to be 40 mg L^{-1} at pH 4.5 and 100 mg L^{-1} at pH 5.0.

Table 5.1 Wet pellet densities recorded for a PAF expressing clone grown in a 7.5 L bioreactor. Culture media was maintained at pH 4.5 or pH 5.0 for 72 hours.

	pH 4.5	pH 5.0
Time (h)	Wet pellet weight (mg/mL)	Wet pellet weight (mg/mL)
Inoculum	46.4	57.1
0	16.4	18.8
24	139.0	87.0
48	206.6	221.0
72	219.8	278.9

A: pH 4.5



B: pH 5.0

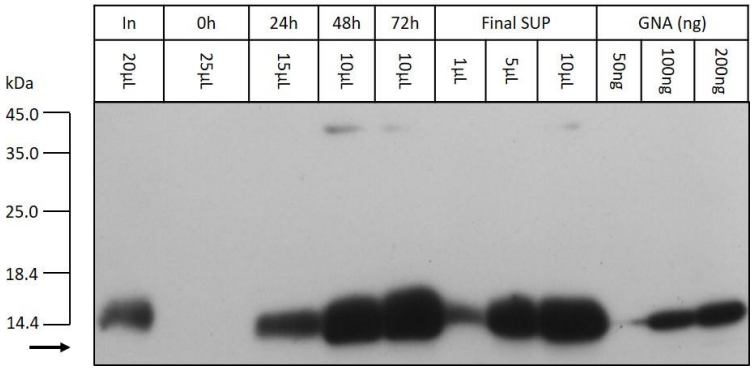


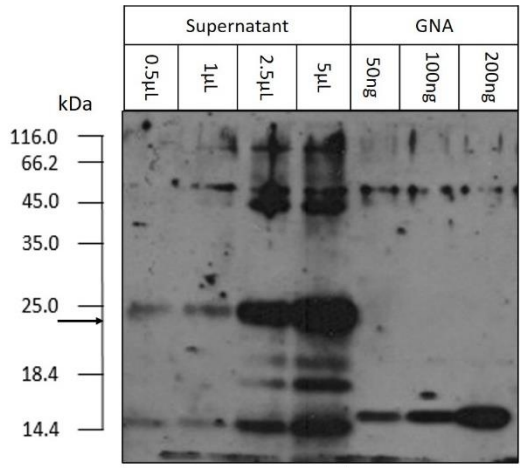
Figure 5.4 Expression level of PAF during fermentation at pH 4.5 (A) and pH 5 (B). Western blot analysis (using anti-His antibodies) of the expression of PAF in fermented culture supernatants over time. A and B are samples from fermentation conducted at pH 4.5 and 5.0, respectively. In denotes inoculum sample and volumes (µl) are stated above each lane. GNA: recombinant GNA standards (quantity in ng stated above lanes). Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

In all samples from PAF/GNA fermentation, immunoreactive bands were present at a mass of approximately 25 kDa, slightly higher than the predicted mass of 23.9 kDa (Figure 5.5). Again, the slightly increased mass of the immunoreactive band compared to the predicted mass may be a result of glycosylation. Each of the supernatants also contained other lower molecular weight immunoreactive bands; these likely correspond to cleavage products of the fusion protein, through differential cleavage of PAF/GNA by yeast proteases. The final supernatant volume of all fermentations was approximately 3 L. Expression levels were estimated semi-quantitively by comparison of the 25 kDa band in the supernatant with recombinant GNA standards of known quantity. When fermented at pH 4.5 with glycerol, PAF/GNA was expressed at 80 mg L^{-1}, compared to 10 mg L^{-1} when fermented at pH 5 (Figure 5.5A and B). Utilising glucose as a carbon source at pH 4.5 caused a decrease in expression level (50 mg L^{-1}) as compared to glycerol (Figure 5.5C).

11z

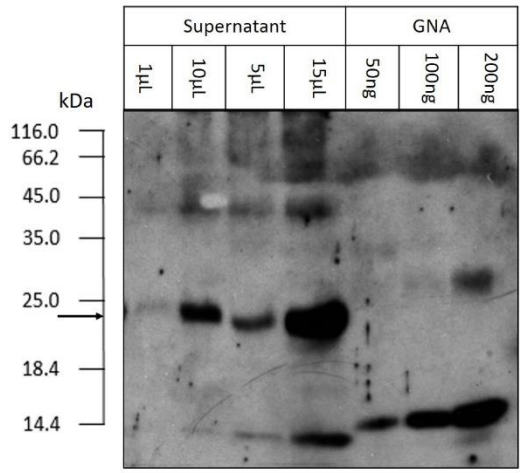
A: pH 4.5;

Carbon Source: Glycerol



B: pH 5.0;

Carbon Source: Glycerol



C: pH 4.5;

Carbon Source: Glucose

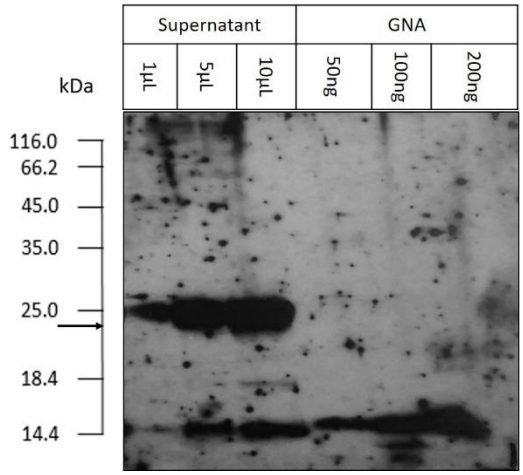


Figure 5.5 Expression of PAF/GNA fermented at pH 4.5 with glycerol (A), pH 5 with glycerol (B) and pH 4.5 with glucose (C). Western blot analysis (using anti-GNA antibodies) of the PAF/GNA in fermented culture supernatants. A, B and C are samples from fermentation conducted at pH 4.5 with glycerol, pH 5.0 with glycerol and pH 4.5 with glucose, respectively. Sample and volumes (µl) are stated above each lane. GNA: recombinant GNA standards (quantity in ng stated above lanes). Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

5.7 Purification of PAF and PAF/GNA from culture supernatant fermented at different pHs with different carbon sources

Recombinant PAF and PAF/GNA from the various fermentations were purified under identical conditions which allows direct comparison of purification efficiency by nickel affinity chromatography and recovery of protein. The supernatant was prepared for purification as described in section 2.8. The combined buffer and supernatant were loaded on to 2 x 5 mL HisTrap FF columns (GE healthcare), for 2 hours at RT, with a flow rate of 5 mL min^{-1}. The final pHs of loaded PAF samples for fermentations conducted at pH 4.5 and pH 5 were pH 4.87 and pH 5.36, respectively. The final loaded pH of PAF/GNA supernatants were pH 4.90 (pH 4.5 and glycerol), pH 5.3 (pH 5 and glycerol) and pH 4.93 (pH 4.5 and glucose). Purification was performed as described in section 2.8. Attempts to raise the pH of the supernatant to 7.4 by the addition of NaOH resulted in precipitation.

Comparison of recombinant protein quantities in supernatant and flow through samples provides an indication of how efficiently recombinant PAF bound to the nickel affinity column. Figure 5.6 showed that there was a marked decrease in immunoreactive band intensity between the supernatant and the flow through, suggesting that recombinant PAF bound efficiently to the nickel columns in both conditions. However, Figure 5.6A shows some of the bound protein is lost in the binding buffer wash (first 5 mL of wash step was collected and used for analysis), presumably due to the increased pH of the wash causing bound protein to be released from the column, which likely reduced the final yield, this band is not visible in Figure 5.6B. In both instances no detectable quantity of protein was lost during the 10 mM imidazole wash step. Comparison of the eluted fractions suggests that greater quantities of PAF were eluted when the sample was loaded at pH 5.37 compared to pH 4.87, though was possibly due to higher concentrations of PAF present in supernatant derived from the pH 5.0 fermentation.

PAF/GNA did not bind efficiently to nickel columns under any of the conditions tested (Figure 5.7). Fermentation at pH 5 with glycerol was the condition under which PAF/GNA bound most efficiently to the nickel columns (Figure 5.7B), indicated by decreased band intensity between the supernatant and flow through samples. However, protein was detectable in the first wash. In all conditions a large intense band corresponding to the predicted weight of PAF/GNA was detected in a small volume of the peak eluted fraction. A 40-45 kDa immunoreactive protein was present in the elution lanes of Figure 5.7B. This could potentially be a dimer, or a PAF/GNA – yeast protein complex. The low molecular weight bands detected in Figure 5.5A were not present in Figure 5.7A. Low molecular weight proteins at 14 kDa were visible in lyophilised samples of both Figure 5.9B and C, suggesting that the loss of mass is due to cleavage of the N-terminus rather than of the C-terminus where the his-tag is located.

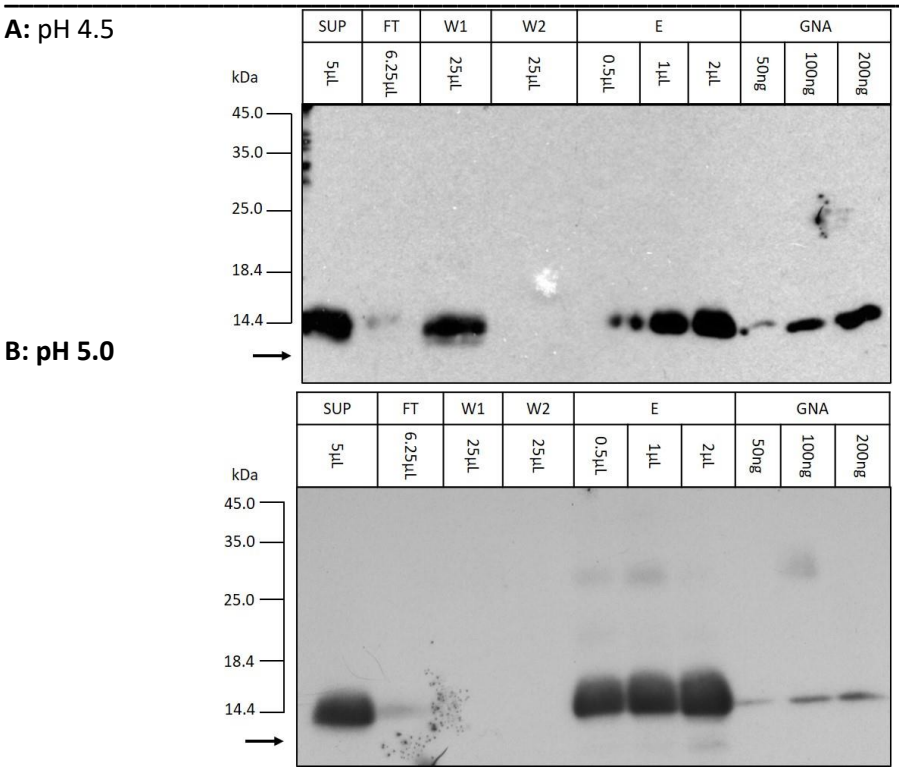


Figure 5.6 Analysis of purification fractions of PAF from fermentation at pH 4.5 (A) and 5 (B). Western blot analysis (using anti-His antibodies) of PAF samples purified by nickel-affinity chromatography. Sample volumes (µl) are stated above each lane. SUP: Supernatant, FT: Column flow through, W1: 1 x binding buffer, W2: 1 x binding buffer + 10 mM imidazole, E: 1 x binding buffer + 200 mM imidazole, GNA: recombinant GNA standards (quantity in ng stated above lanes). Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

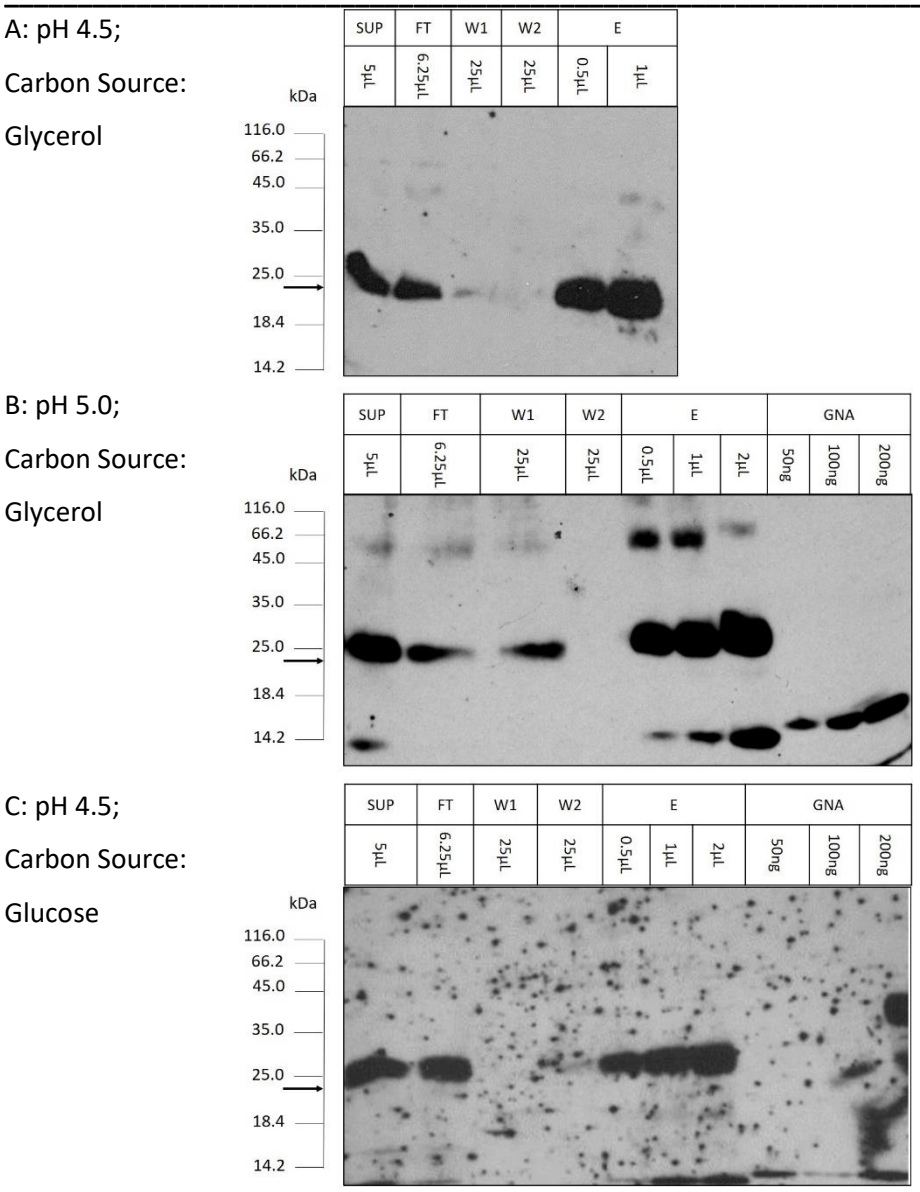


Figure 5.7 Analysis of purification fractions of PAF/GNA from fermentation at pH 4.5 with glycerol (A), pH 5.0 with glycerol (B) and pH 4.5 with glucose (C). Western blot analysis (using anti-GNA antibodies) of PAF/GNA samples purified by nickel-affinity chromatography. Sample volumes (µl) are stated above each lane. SUP: Supernatant, FT: Column flow through, W1: 1 x binding buffer, W2: 1 x binding buffer + 10 mM imidazole, E: 1 x binding buffer + 200 mM imidazole, GNA: recombinant GNA standards (quantity in ng stated above lanes). Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side.

5.8 Quantification and recovery of recombinant PAF and PAF/GNA

After purification, the eluted fractions were dialysed and lyophilised as described in section 2.10. The protein content of the lyophilised powder was estimated semi-quantitively by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) stained with coomassie blue by visual comparison of known quantities of lyophilised powder to commercially available GNA standards (Sigma-Aldrich, St. Louis, USA).

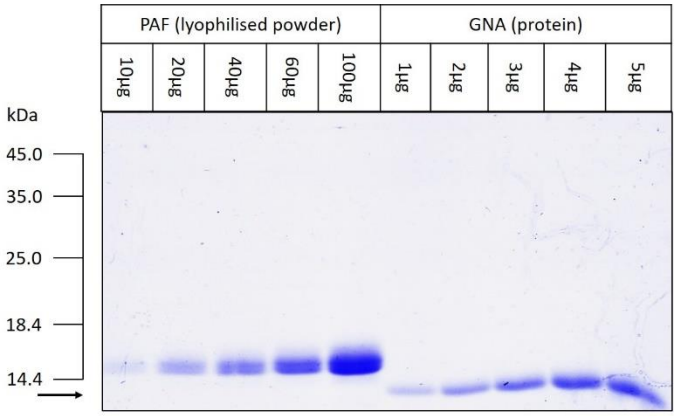
Representative quantification analysis for PAF samples is shown in Figure 5.8. All samples contain a protein band of approximately 14 kDa, which corresponds to the mass observed in the western blots but is greater than the predicted MW of PAF. The proportion of PAF in powder derived from fermentation at pH 4.5 was 10 % (w/w). When fermented at pH 5, the percentage of PAF in lyophilised powder was significantly higher, at 50 % (w/w). Purification from cultures fermented at pH 4.5 recovered 8 mg, of an estimated 40 mg present in 1 L of supernatant, equating to a 20 % protein recovery. Recovery of PAF from the pH 5.0 supernatant was much greater; 42 mg of protein was recovered from 1 L of supernatant estimated to contain approximately 100 mg PAF resulting in a 42 % protein recovery. The protein which is unaccounted for could remain in the supernatant, as a faint immunoreactive band was detected in the flow throughs (Figure 5.6). Or perhaps is insoluble after lyophilisation.

Quantification of PAF/GNA is shown in Figure 5.9. All samples contained a protein of approximately 25 kDa, corresponding to the immunoreactive bands detected by western blotting. The different fermentation conditions resulted in different profiles on gel. Part A of Figure 5.9 has two additional major bands, at approximately 21 kDa and 14 kDa which likely correspond to cleavage products of PAF/GNA. Based on the molecular weight and immunoreactivity to anti-GNA antibodies, these bands are probably PA1a/GNA (21 kDa), where the PA1b toxin has been cleaved, and where the entire PAF domain has been cleaved from GNA (14 kDa) (Figure 5.5A, Figure 5.9A). Two proteins of approximately 21 kDa and 14 kDa are also present in the coomassie stained gel (Figure 5.9B); however, they are much fainter and only the 14 kDa band is immunoreactive to anti-GNA antibodies (Figure 5.9B). The proportion of PAF/GNA in the lyophilised powders from fermentation at pH 4.5 with glycerol, pH 5.0 with glycerol and pH 4.5 with glucose are 10 % (w/w), 15 % (w/w) and 8.75 % (w/w), respectively. These lyophilised powder concentrations were generally poor. The total protein recovered from the supernatant and percentage recovery for both PAF and PAF/GNA fermentations are displayed in Table 5.2. The low percentage of protein recovered in the pH 4.5 fermentations can be partially attributed a proportion of the protein being insoluble after lyophilisation. Figure 5.10 shows both PAF and PAF/GNA which has been resuspended in PBS, and then the pellet resulting from that resuspension was subsequently resuspended in SDS loading buffer. In all pH 4.5 conditions insoluble protein was present at the appropriate mass.

Table 5.2 Summary of data pertaining to the production and purification of PAF and PAF/GNA. Data are based on purification of 1 litre of supernatant.

	PAF		PAF/GNA		
pH	4.5	5	4.5	5	4.5
Carbon source	Glycerol	Glycerol	Glycerol	Glycerol	Glucose
Expression level (mg L^{-1})	40	100	80	10	50
Lyophilised powder (% protein w/w)	10	50	10	15	8.75
Protein recovered (mg L^{-1})	8	42	5.8	3.5	18.2
Recovery (%)	20	42	7	35	36

A: pH 4.5



B: pH 5.0

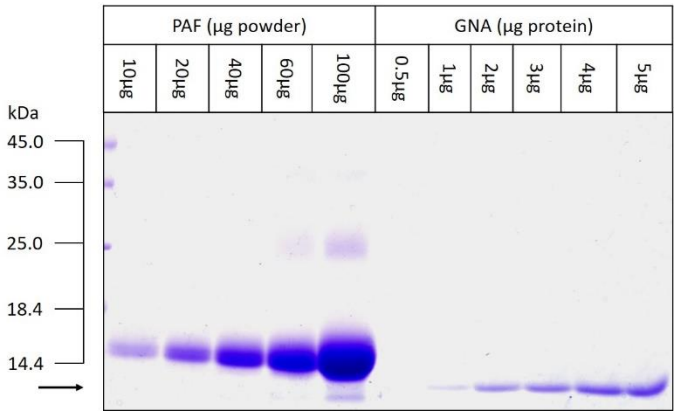


Figure 5.8 Quantification of PAF in lyophilised powder derived from fermentation at pH 4.5 (A) and pH 5 (B). Coomassie brilliant blue stained SDS-PAGE gels (17.5% acrylamide) of recombinant PAF after purification by nickel affinity chromatography, lyophilisation and resuspension in PBS. A and B are samples from fermentation conducted at pH 4.5 and 5.0, respectively. Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

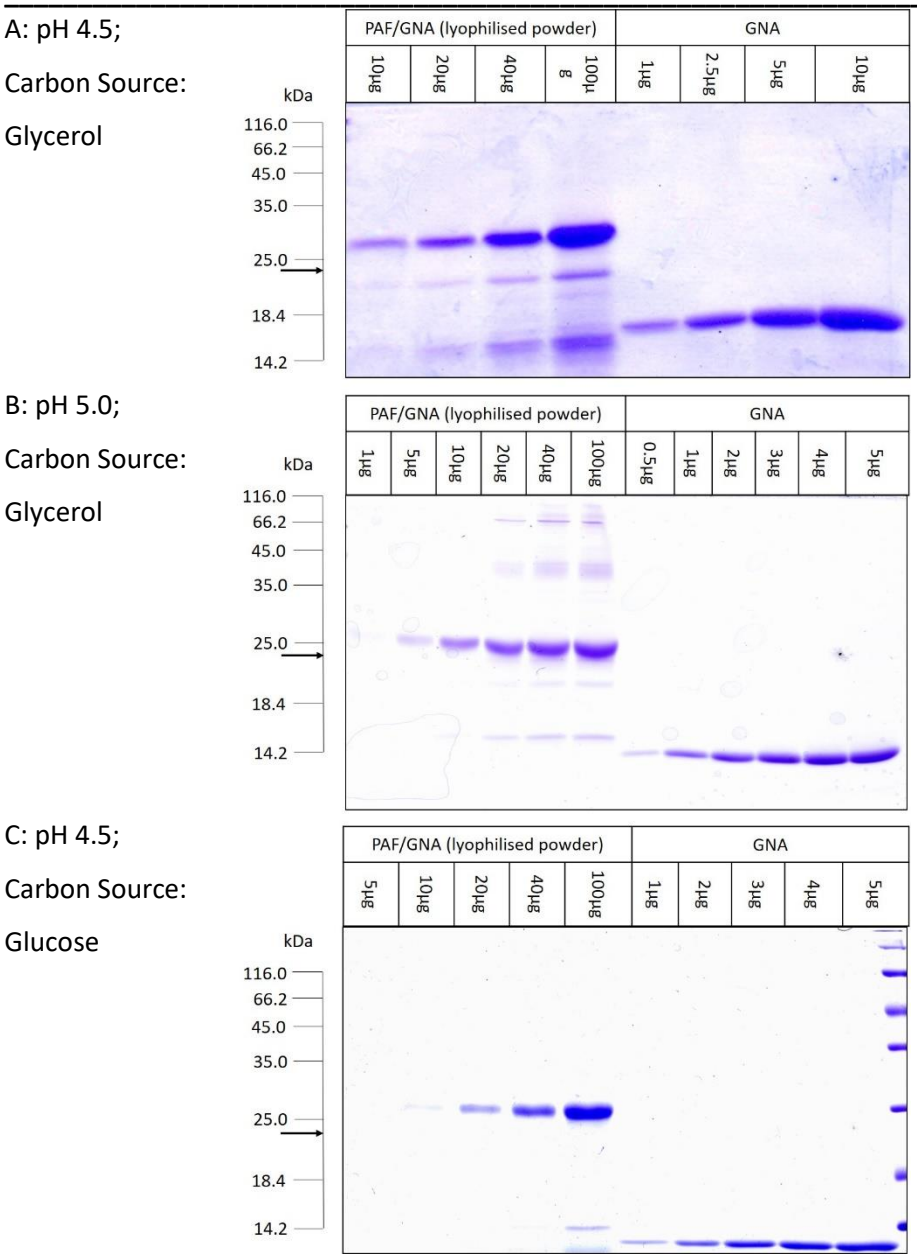


Figure 5.9 Quantification of PAF/GNA in lyophilised powder derived from fermentation at pH 4.5 with glycerol (A), pH 5 with glycerol (B) and pH 4.5 with glucose (C). Coomassie brilliant blue stained SDS-PAGE gels (17.5% acrylamide) of PAF/GNA after purification by nickel-affinity chromatography, lyophilisation and resuspension in PBS. GNA standards were commercial purified GNA from snowdrop bulbs (Sigma-Aldrich, St. Louis, USA). Quantities of powder and protein loaded are stated above the lanes in µg. Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

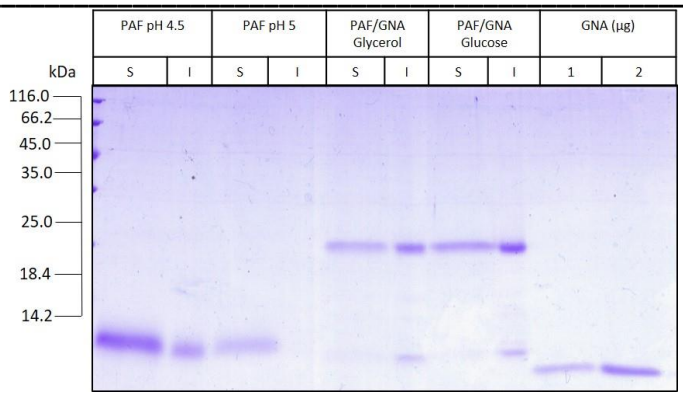


Figure 5.10 Soluble and insoluble fractions of lyophilised powder of PAF and PAF/GNA. Lyophilised powder with protein content of 5 ug was dissolved in 1 x PBS and designated soluble. The resulting pellet was then dissolved in 1 X SDS loading buffer and designated insoluble. S: 2 µg of the soluble fraction, I: equivalent volume to 2 µg of the soluble fraction. Sainted with Coomassie brilliant blue. Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side.

5.8 Peptide identification by LC-MS

Recombinant protein from the best performing fermentations (PAF pH 5.0 and PAF/GNA pH 4.5 with glucose) were sent for analysis by liquid chromatography mass-spectrometry. Marked bands were excised from an SDS-PAGE acrylamide gel and trypsin digested, then analysed by LC-MS as described in section 2.14. Formic acid blank injections were performed between each sample injection.

PAF runs as a single intact band at approximately 14 kDa on SDS-PAGE gels. Analysis by LC-MS revealed that both the PA1b and PA1a peptides are present in the band, indicating the protein is intact. The detected peptide fragments are shown in Figure 5.11. The α-factor was correctly cleaved, apart from five additional N-terminal residues (Glu-Ala-Glu-Ala-Ala) preceeding PAF. These additional residues partially explain the increased mass of PAF when analysed by SDS-PAGE. A sequence of peptides in the centre of PAF was not detected but this is likely to be due to a long sequence of residues without arginine or lysine residues to facilitate cleavage by trypsin.

Two recombinant PAF/GNA protein products, one at 25 kDa and one at approximately 14 kDa, (Figure 5.12) when analysed by LC-MS confirmed the presence of PAF/GNA in the 25 kDa band. As for PAF, five additional N-terminal residues (Glu-Ala-Glu-Ala-Ala) were detected in the 25 kDa PAF/GNA protein products, providing a partial explanation for the increased mass on SDS-PAGE compared to the predicted mass of PAF/GNA with the additional residues adding approximately 0.5 kDa. A 25-peptide sequence at the N-terminus of GNA was not detected, possibly due to the large distance between lysine or arginine residues. The 14 kDa band is likely to be a cleavage product of PAF/GNA, with no spectra matching to the N-terminus of PAF and limited spectra matching to N- and C-terminals of GNA (Figure 5.12). The spectra detected between residue 113 and 148 is likely carry over from the previous sample injection, as this peptide was detected in the formic acid blank and in the GNA only sample, which was injected last, in addition to this the mass from the sequence STSVAAAA to the end of the his-tag is 14 kDa suggesting that the 14 kDa band is a GNA cleavage product of PAF/GNA (Figure 5.13 and Figure 5.14). Correct cleavage of the alpha factor from GNA could not be determined as LC-MS analysis failed to detect the first 30 peptides of the protein. The contaminating peptide from PAF/GNA is visible in the GNA only sample, suggesting it has been carried over from the first injection, through multiple injections and blanks (Figure 5.14).

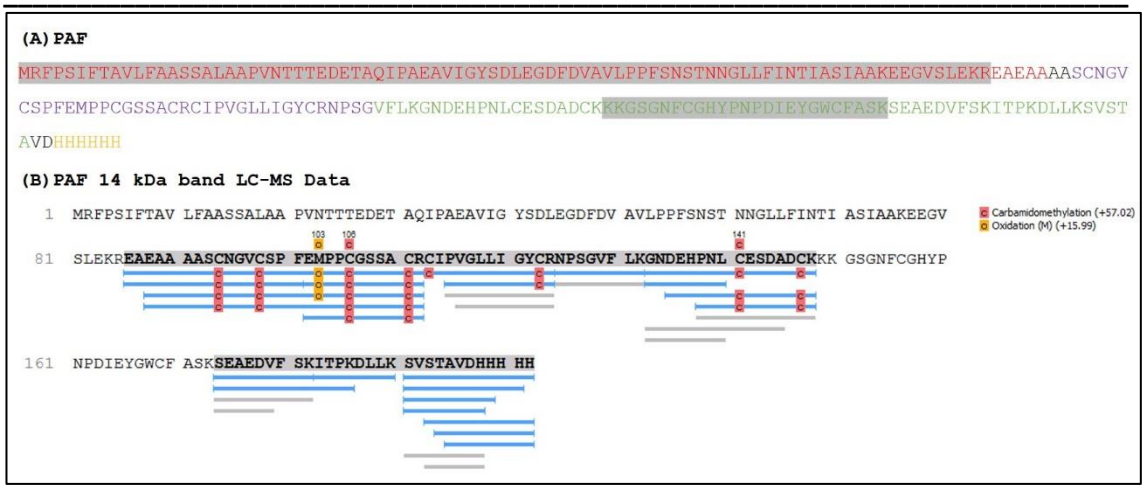


Figure 5.11 Primary structure and LC-MS data from the PAF 14 kDa band on SDS-PAGE. A: Primary structure of the recombinant PAF protein product expressed by transformed *P. pastoris* cells. The alpha factor secretory signal is depicted in red, PA1b in purple, PA1a in green, and the his-tag in yellow. Remaining residues are the result of the cloning process for the expression construct. Residues highlighted in grey were not detected in the 14 kDa band see (B). B: Peptide matches of the 14 kDa band compared to the expressed sequence.

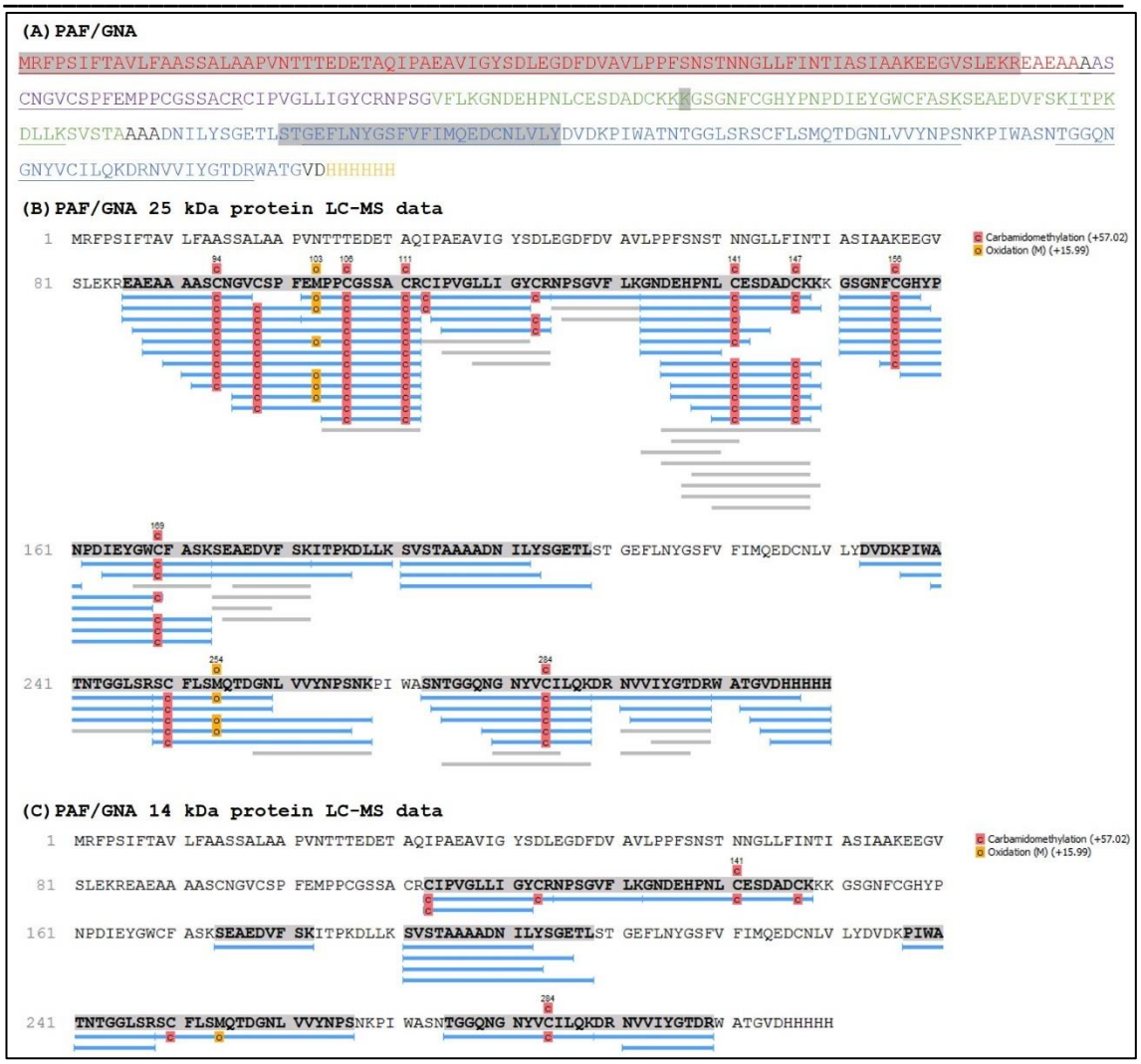


Figure 5.12 Primary structure and LC-MS data from the PAF/GNA 25 kDa and 14 kDa bands on SDS-PAGE. A: Primary structure of the recombinant protein expressed by transformed *P. pastoris* cells. The alpha factor secretory signal is depicted in red, PA1b in purple, PA1a in green, GNA in blue and the his-tag in yellow. Remaining residues are the result of the cloning process for the expression construct. Residues highlighted in grey were not detected in the 25 kDa band (see B). Underlined residues were not detected in the 14 kDa band (see C). B: Peptide matches of the 25 kDa band compared to the expressed sequence. C: Peptide matches of the 14 kDa band compared to the expressed sequence.

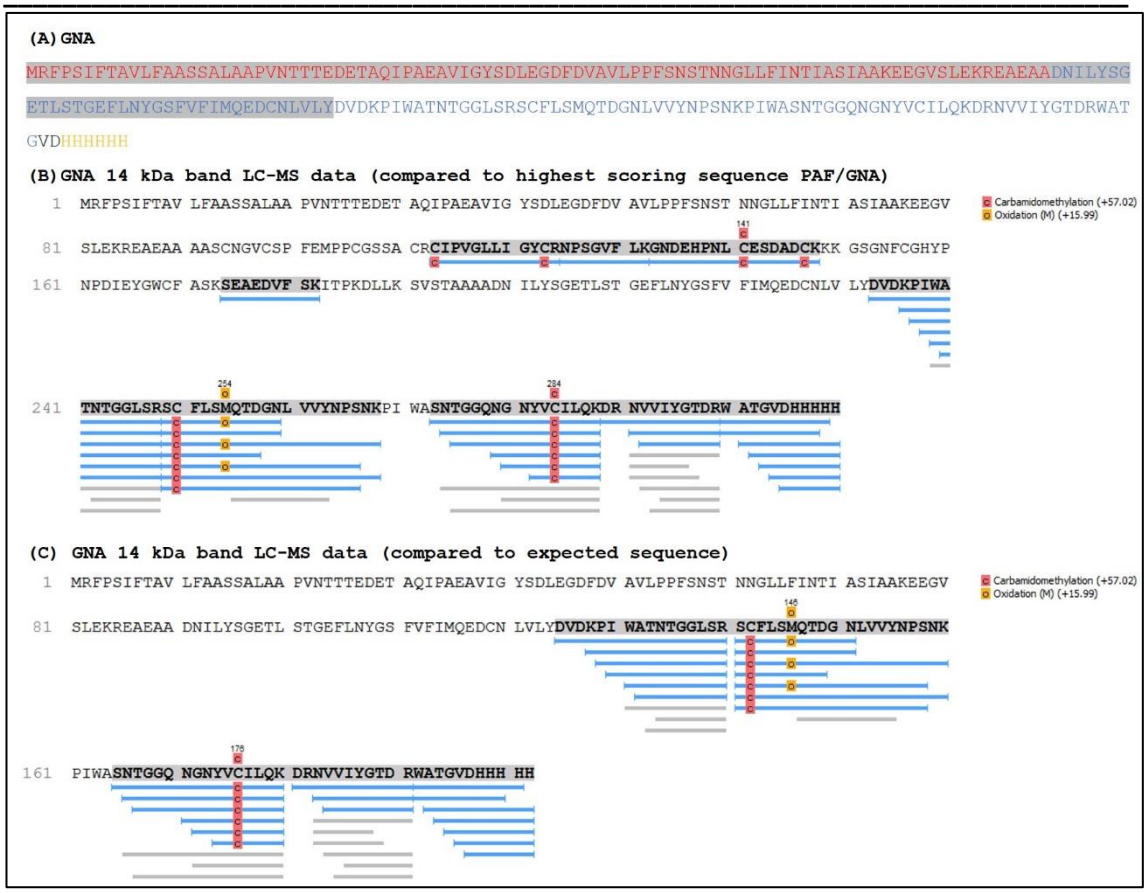


Figure 5.13 Primary structure and LC-MS data from the GNA 14 kDa band on SDS-PAGE. A: Primary structure of the recombinant protein expressed by transformed *P. pastoris* cells. The alpha factor secretory signal is depicted in red, GNA in blue and the his-tag in yellow. Remaining residues are the result of the cloning process for the expression construct. Residues highlighted in grey were not detected in th 25 kDa band see (B). B: Peptide matches of the 14 kDa band compared to the top scoring sequence (PAF/GNA). C: Peptide matches of the 14 kDa band compared to the expected sequence.

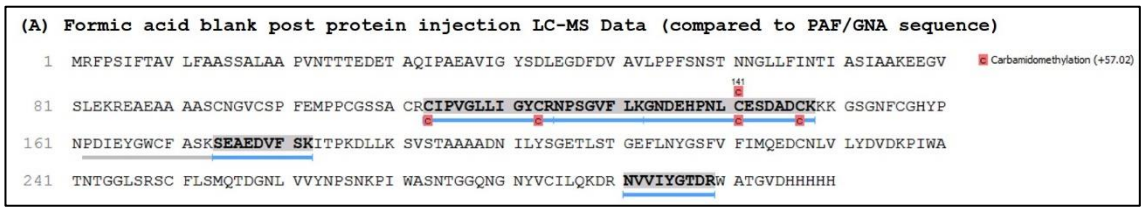


Figure 5.14 LC-MS data from the formic acid blank. Data are compared to the sequence for PAF/GNA.

5.9 Glycosylation of PAF and PAF/GNA fermented under different pH and carbon sources

Pichia pastoris is known to glycosylate recombinant proteins during expression (Macauley-Patrick *et al*., 2005). There is one potential N-linked site (N-P-S) present in both PAF and PAF/GNA, beginning at the 34th amino acid residue of the PAF peptide, which fits the required structure described in Schwarz and Aebi (2011) (Figure 5.1, Figure 5.2). The glycosylation of recombinant PAF and PAF/GNA from both fermentations was analysed by Schiff-periodic acid blotting. Visualisation of the protein bands with Schiff reagent showed glycosylation of PAF from both fermentations as a pink band of approximately 14 kDa. None of the three samples of PAF/GNA showed a pink band at 25 kDa, where the protein band is visible in the Coomassie stained gel suggesting that this protein product is not glycosylated (Figure 5.15). Staining was observed at high molecular weights in the PAF/GNA lanes, suggesting minor contamination with highly glycosylated, high molecular weight protein, which is not present in high enough concentrations to be visible by Coomassie staining. Schiff reagent stains glycoproteins in a quantitative manner; therefore, it appears that the changes undertaken in fermentation method have little effect on the quantity of glycosylation of the final recombinant proteins. Large quantities of glycosylation are observed in the positive control, which is PI1a/GNA, a fusion protein that is known to be glycosylated (N-linked) (Yang *et al*., 2014). The negative control was commercially available purified GNA, purified from snowdrop bulbs (Sigma-Aldrich, St. Louis, USA) which is not glycosylated (Van Damme *et al*., 1987).

5.10 Hemagglutination assays (PAF/GNA)

The biological activity of the GNA component of PAF/GNA was assessed using hemagglutination assays as described in section 2.12. The method through which hemagglutination functions with respect to GNA is described in section 2.12. Agglutination experiments were performed with lyophilised powders derived from both pH 4.5 and pH 5 media that used glycerol as a carbon source. As can be seen in Figure 5.16, when fermented at pH 4.5 the resulting protein could agglutinate rabbit erythrocytes to a concentration of 7.8 µg mL^{-1}, whereas the protein derived from fermentation at pH 5 only agglutinated to a concentration of 125 µg mL^{-1}. This is a decrease in function of the GNA lectin and could be due to issues with protein folding when produced at higher pHs.

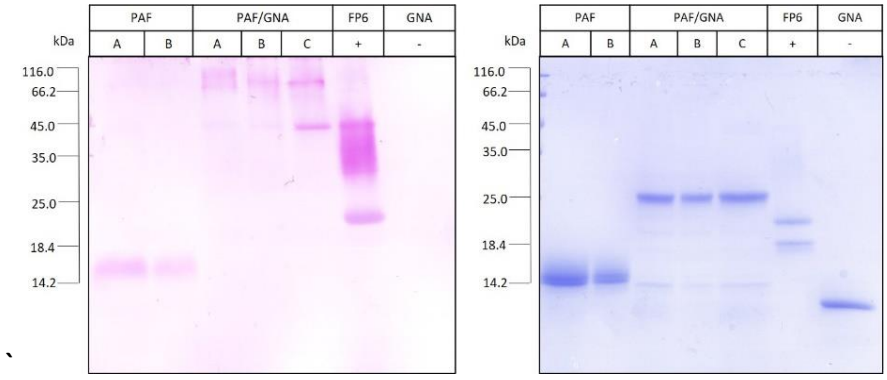


Figure 5.15 Schiff blot to detect glycosylation of PAF and PAF/GNA from different fermentations. Left: Schiff Blot 10 ug protein loaded in each lane. Right: Coomassie Stained poly acrylamide gel, 5 ug loaded in each lane. PAF A: fermented at pH 4.5, B: fermented at pH 5.0. PAF/GNA A: fermented at pH 4.5 with glycerol, B: fermented at pH 5.0 with glycerol, C: fermented at pH 4.5 with glucose. FP6 was included as a positive control. GNA: Commercially purified GNA (Sigma-Aldrich, St. Louis, USA). Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side.

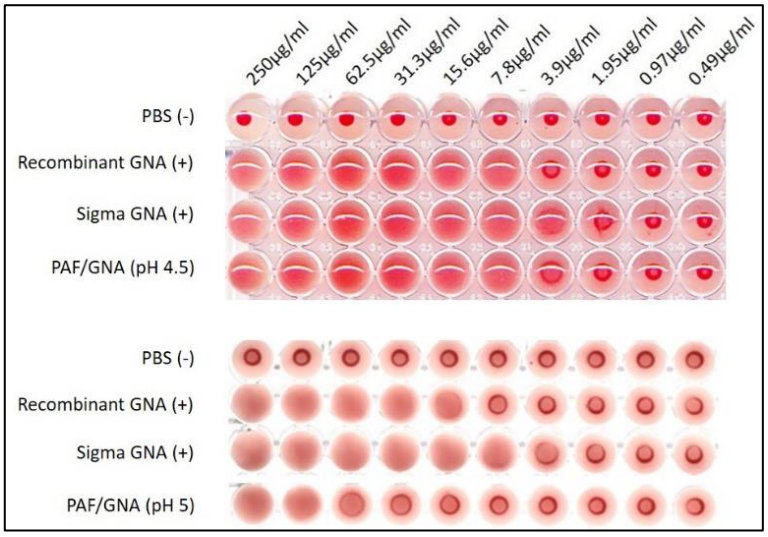


Figure 5.16 Agglutination assay to detect function of GNA when fused to the C-terminus of PAF. Protein was serially diluted and incubated with a 2% (v/v) rabbit erythrocyte solution. After 2 hours the plate was imaged to determine the lowest concentration at which GNA can crosslink rabbit erythrocytes. PBS was used as a negative control and recombinant GNA and Commercial (Sigma-Aldrich, St. Louis, USA) GNA were included as positive controls. Agglutinated or crosslinked cells display an even red colour in the well whilst non-agglutinated cells settle to the bottom of the plate and form a visible dark red pellet.

5.11 Discussion

Many factors can significantly alter the production of recombinant proteins in *P. pastoris,* such as pH, cell density, *P. pastoris* strain, promoter, temperature, dissolved oxygen percentage and carbon source (Li *et al*., 2007). Specifically, this study focused on pH and carbon source as factors as one of the main hypotheses as to why binding was poor was related to pH of the culture media, and previous work (Chapter 4) demonstrated that using glucose in the medium can lead to increased protein recovery compared to glycerol. *Pichia pastoris* was selected as the expression system for PAF and PAF/GNA as it is well suited for growth in bioreactors and due to its ability to produce post-translational modifications typically associated with protein produced by higher eukaryotes (Cregg *et al*., 2000). The ability to produce disulphide bridges is key as PA1b is a small inhibitor cysteine knot protein, as such it contains three disulphide bridges (Jouvensal *et al*., 2003). These disulphide bonds are vital in determining the structure and function of PA1b as they maintain the loop structures which contain the target binding sites (Da Silva *et al*., 2010). A complete version of the PA1 gene cassette (PA1b + PA1a) was used in this chapter, as issues were encountered when trying to express PA1b on its own in the literature, in previous work performed in our lab and when expressing and purifying PA1b/GNA (discussed in chapter 4) (Eyraud *et al*., 2013), Pyati - personal comm 2016). This chapter focuses on the optimisation of fermentation conditions to produce recombinant pea albumin (PAF) and pea albumin – GNA fusion proteins (PAF/GNA) in BSM. Specifically, it focused on the effects of pH and (due to the results of Chapter 4), the effect of glucose was tested on PAF/GNA. Evidence is presented to suggest that the expression and purification of PAF is significantly improved when pH was maintained at 5 as opposed to 4.5 during bench-top fermentation. Furthermore, fermentation of PAF/GNA with glucose at pH 4.5 was determined to be optimal as, despite the reduced expression level, the quantity of final protein recovered was greater. The proteins from the best performing fermentations were analysed by LC-MS to confirm correct expression and processing.

Increasing the pH of the fermentation from pH 4.5 to pH 5 did not appear to significantly impact the growth of *P. pastoris*. The initial inoculum pellet weight for the pH 5 fermentation was 1.2 fold greater than that for the pH 4.5 run, and the same difference in pellet weight is still present at the end of the fermentations implying that PAF expressing *P. pastoris* is growing at comparable rates in both pH 4.5 and pH 5 media. This is also evidenced when the specific growth rate was calculated 0.036 μ h^{-1} (pH 4.5) compared to 0.037 μ h^{-1} (pH 5.0). These growth rates were comparable to that observed in Chapter 3, and greater than observed by Prielhofer *et al*., (2015) but this is likely due to methodological differences, in particular shake flask versus

bioreactor and differences in media. Differences cell density (1.2-fold greater in pH 5) or growth rate is therefore unlikely to completely account for the difference in expression level observed between pHs (2.5-fold increase between pH 4.5 and 5).

Optimal pH for expression of recombinant proteins in *P. pastoris* cover a broad range from pH 3.0 to 7.0 (Çalık *et al*., 2010; Chang *et al*., 2006; Eissazadeh *et al*., 2017; Koganesawa *et al*., 2002). Other small insecticidal knottin proteins and fusion proteins expressed *in P. pastoris* were generally fermented between pH 4.5 and 5.0, which is why these two pHs were selected (Fitches *et al*., 2004, 2012; Pyati *et al*., 2014). This study did not test pHs above 5 due to increased salt levels and mineral precipitation in the BSM observed when pH of the media is raised above 5. Alternate growth media, such as YPG and BMGY, which are normally used for neutral and high pH fermentations of *P. pastoris,* were not tested due to the results from in Chapter 4 where both alternate media resulted in the presence of many contaminating proteins in the final lyophilised powder. The estimated expression level of PAF was found to be higher by a factor of 2.5 when fermented at pH 5 compared to pH 4.5. Prediction of optimal fermentation conditions is difficult; *Pisum sativum* defensin (47 aa, eight cysteine residues) was shown to have the greatest expression at pH 6 (Almeida *et al*., 2001); whereas human growth hormone (191 aa, 4 cysteines) displayed optimal expression at pH 5, similar to PAF, despite its difference in size and structure (Çalık *et al*., 2010) The increased expression level cannot be explained entirely by the difference in cell density as described above. Jahic *et al*., (2003) reported that proteolysis of their fusion protein decreased when the pH of their bioreactor culture was lowered from pH 5.0 to pH 4.0. This in combination with the use of a protease deficient strain of *P. pastoris* (SMD1168H) suggests that decreased protease activity is not responsible for the increased quantities of protein in the pH 5.0 fermentation. The iso-electric point of PAF was predicted to be 6.33 so precipitation during fermentation due to pH proximity to its pI is unlikely in either condition (https://web.expasy.org/compute_pi/). Perhaps PAF folded into the correct confirmation more efficiently at the higher pH, which is likely closer to the *in vivo* pH at which PA1b is expressed *in planta*.

Conversely the estimated expression level of PAF/GNA was much higher when fermented at pH 4.5 (80 mg L^{-1}), as opposed to pH 5 (10 mg L^{-1}), exemplifying the fact that optimal conditions for protein production cannot be predicted and must be experimentally determined on a case by case basis. Fermentation with glucose was performed at pH 4.5, resulting in a somewhat decreased expression level of 50 mg L^{-1} compared to the glycerol fed fermentation. The indicated decrease in expression may be attributable to cell density but as wet pellet weights were not collected it is not possible to comment on this. The values obtained

for the pH 4.5 fermentations are consistent with reported values for large scale expression of recombinant GNA by *P. pastoris* (Baumgartner *et al*., 2003). Decreased expression level in glucose-fed fermentations is contrary to published literature for proteins expressed under the control of the GAP promoter but is consistent with results in Chapter 3 (Waterham *et al*., 1997).

The variability of expression level observed between different fermentations is not surprising. Quantities of protein produced per litre in bioreactors vary wildly in the literature, proteins such as gelatins and human serum albumin both were reported protein content levels in the broth of 10 g L^{-1} or greater, whilst other proteins as little as 1.0 mg L^{-1} were reported, a selection of proteins are summarised in Macauley-Patrick *et al*., (2005) (Sumi *et al*., 1999; Werten *et al*., 1999). The expression levels reported in this study are on the lower end of the spectrum but fermentation at more optimal pH was shown to positively affect the expression level of both proteins.

Purification by nickel affinity chromatography relies on the interaction of the two histidine side chains with a positively charged Ni^{2+} ion bound to a resin (Bornhorst and Falke, 2000). PAF bound efficiently to the columns under both fermentation conditions, evidenced by the lack of immunoreactive protein in the flow through when analysed by western blot. An immunoreactive protein was present in the 1 x binding buffer wash of the pH 4.5 condition, suggesting that the protein was not bound tightly to the column and was lost during the washing step. This provides an obvious reason for the difference in recovery observed between the two fermentations (pH 4.5: 20 %, pH 5.0: 42 %). The pH in the column would increase rapidly (4.87 to 7.40), through the pI of the protein, when washing with binding buffer. This rapid change in pH may be responsible for the lost protein. At the loading pH of 4.87 the histidine residues will have been in a protonated state, as the pH is below the pI of the protein (6.33), which reduces histidine's capacity to interact with nickel (Bornhorst and Falke, 2000).

Little decrease was observed in the intensity of immunoreactive bands in flowthough and the supernatant samples derived from any of the PAF/GNA fermentations. Poor protein column interaction could also be attributed to his-tag availability as was observed with PA1b/GNA in Chapter 4. Early attempts to purify PA1b/GNA under denaturing conditions failed to result in the recovery of any soluble protein (Chapter 4); whilst PA1b/GNA bound more efficiently under denaturing conditions, precipitation during re-folding dialysis using a urea gradient resulted in negligible recovery of soluble protein. Therefore, denaturing purification was not attempted with PAF/GNA. A second hypothesis for poor binding efficiency is that as the pH of the purification buffer was below the pI of the protein (5.54) the his-tag was protonated, thereby resulting in reduced binding affinity (Bornhorst and Falke, 2000). As stated previously

attempts to raise the pI of the protein resulted in precipitation when the pH neared the pI of the protein. Poor column interaction is suspected to be the major reason for low percentage recovery of this protein based on the data presented. PAF/GNA fermented at pH 5.0 had the most efficient interaction with nickel, showing the largest decrease in intensity between supernatant and flow through fractions. However, it also was the only condition in which protein was lost during the binding buffer wash, suggesting poor binding of the protein to the purification column.

Soluble PAF concentration in the lyophilised powder derived from pH 4.5 fermentation was 10 % (w/w) and from pH 5 was 50 % (w/w). Greater percentages of protein in the lyophilised powder is preferred, as it suggests reduced quantities of carbohydrate or insoluble protein, which as PAF is glycosylated makes up the remaining powder. Total PAF/GNA recovered (42 mg versus 8 mg) and percentage recovery was improved in the pH 5 fermentation (42 % versus 20 %), possibly as the higher loading pH allowed more protein to bind to the nickel matrix. One hundred percent recovery is uncommon during protein purification as protein can be lost at each step; protein may not bind to the columns and remain in the supernatant, may precipitate during dialysis, or may be insoluble after lyophilisation, evidence for the latter was observed in both PAF pH 4.5 and PAF/GNA pH 4.5 fermentations. Analysis of the major recombinant protein product by LC-MS suggests that PAF is expressed intact and includes five additional N-terminal residues (Glu-Ala-Glu-Ala-Ala, 0.5 kDa) that were not cleaved by the yeast Ste13 gene product (Dipeptidyl aminopeptidase A). Additional N-terminal residues in recombinant proteins secreted using the *S. cerevisiae* α-factor secretory signal have been observed previously. Analysis of Hv1a/GNA, GNA/Hv1a and another sulphur rich *P. sativum* protein by N-terminal sequencing and LC-MS detected the presence of these amino acids (Almeida *et al*., 2001; Powell *et al*., 2019). The presence of these residues and the glycosylation detected by the Schiff assay provide an explanation for the increased mass of PAF observed on gel, compared to its predicted sequence.

Variation was observed in the quantification and recovery of PAF/GNA. Multiple protein products were detected by SDS-PAGE after lyophilisation of purified samples, suggesting that this protein, as with other insecticidal fusion proteins, is prone to cleavage by yeast proteases under these conditions, even when expressed in protease deficient yeast (Fitches *et al*., 2004, 2012; Trung *et al*., 2006). Based on the mass of these bands, it is suspected that the band at 21 kDa observed in samples derived from the pH 4.5 glycerol fermentation is likely a cleavage product of PAF/GNA where the PA1b peptide has been processed as it would be *in vivo* leaving PA1a/GNA (Higgins *et al*., 1986). The lower 14 kDa bands are likely cleaved GNA products; this is often observed in toxin-GNA fusion proteins (Powell *et al*., 2019; Pyati *et al*., 2014). Only the

25 kDa and 14 kDa band was visible in the pH 4.5 glucose fermentation. Based on the LC-MS analysis of the bands present in the glucose sample, it is evident that 25 kDa band represents the intact fusion protein. The N-terminal region of GNA was not detected, but it was also not detected in the GNA only sample and a similar pattern was observed by Powell *et al*., (2019). The lack of detection was attributed to a large region without trypsin cleavage sites. Similarly, to PAF, the extra five residues were also detected on the N-terminus of PAF/GNA, which provides an explanation for the increased mass of PAF/GNA observed on gel compared to the predicted mass, particularly as no glycosylation was detected by the Schiff blot. The lower band was identified as a cleavage product of PAF/GNA despite containing a fragment of PA1b as this peptide was detected in the blank between sample injections and in the GNA only sample. PAF/GNA fermented at pH 5 displayed reduced agglutination ability when compared to PAF/GNA fermented at pH 4.5. Although this was analysed in separate experiments, the controls were comparable. The agglutination threshold for pH 5 PAF/GNA was 10 times higher than that of pH 4.5 PAF/GNA, indicating that the GNA portion of PAF/GNA derived from fermentation at pH 5.0 is not as functional as that of PAF/GNA derived from pH 4.5 fermentation. The deference in activity of GNA observed may be due to formation of the tertiary structure of GNA; fermentation at higher pH may result in incorrect folding of the final protein product when fused to PAF. Based on this data PAF/GNA from the pH 5.0 fermentation was not used in subsequent bioassays (Chapter 6).

Based on evidence discussed in this chapter, it is apparent that of the conditions tested, the optimal fermentation pH for the expression and recovery of recombinant PAF is pH 5.0. By contrast, PAF/GNA was produced optimally at pH 4.5 using glucose as a carbon source. Although a lower level of expression was observed for PAF/GNA fed with glucose as compared to glycerol, a greater quantity of purified recombinant protein was recovered by affinity chromatography supporting the findings of Chapter 3. It is difficult to speculate on the specific reasons for why these conditions were optimal, as there are many factors which significantly effect protein production and even similar proteins can have different optimal conditions. The results presented here highlight the importance influence that differing pH and carbon sources can have on the recovery of intact and functional recombinant proteins.

Chapter 6 |Biological activity of recombinant pea albumin (PAF) and PAF/GNA

6.1 Introduction

Pea albumin, PA1b, is a small inhibitor cysteine knot protein, isolated from the seeds of the garden pea (*Pisum sativum*). PA1b purified from pea flour using HPLC, has been shown to be orally toxic towards a number of different insect species through specific inhibition of V-ATPase proton pumps in the insect gut epithelium (Eyraud *et al*., 2013; Gressent *et al*., 2007, 2011; Higgins *et al*., 1986) PA1b is currently the only known proteinaceous inhibitor of V-ATPase proton pumps. It selectively inhibits some insect V-ATPases but does not interact with mammalian V-ATPases (Chouabe *et al*., 2011; Gressent *et al*., 2007; Muench *et al*., 2014; Rahioui *et al*., 2014). V-ATPase proton pumps are vital for proton gradient generation which facilitate numerous transport processes and compartment acidification (Beyenbach and Wieczorek, 2006; O'Donnell, 2017). PA1b has also been shown to induce apoptosis in *Sitophilus oryzae* strain WAA42 when delivered orally (Eyraud *et al*., 2017). The existing oral activity to this peptide towards some pest species, with no known effect in mammals, makes it potentially ideal for development as a novel insecticidal protein (Rahioui *et al*., 2014). The lectin *Galanthus nivalis* agglutinin (GNA) has been shown to bind luminal mannose residues in the in the homopteran gut (Powell *et al*., 1998). The toxicity of PAF may be enhanced by linking it to GNA, which could act as an anchor increasing the local concentration of PA1b at the gut epithelium and allowing retention near its target site, the luminal side of the V-ATPase transmembrane domain (Muench *et al*., 2014). Linking other small inhibitor cysteine knot proteins to GNA has been shown to enhance their oral activity by facilitating transport across the gut epithelium and thus access to the central nervous system. This may allow PAF to do the same as V-ATPases are ubiquitous in many eukaryotic membranes, although this is not investigated within the scope thesis (Fitches *et al*., 2002, 2004, 2012; Forgac, 2007). Due to problems with the production and purification of PA1b/GNA (discussed in Chapter 4), this chapter focuses on the recombinant version of the PAF protein (PA1b+PA1a) and the fusion protein PAF/GNA (Chapter 5). The method and optimisation of the production and purification of the proteins tested in this chapter are discussed in Chapter 5. This chapter focuses on testing the insecticidal activity and mechanisms responsible as well as focusing on potential mechanisms for differential toxcictiy observed across different species.

6.2 Oral insecticidal activity of recombinant PAF, GNA and PAF/GNA against aphids

Recombinant PAF, GNA or PAF/GNA were fed to nymphs of *Acrythosiphon pisum, Myzus persicae* or *Sitobion avenae* by incorporating recombinant protein dissolved in PBS or SPB into the artificial diet as described in section 2.17. The control diet was made up to the same volume as the test diet with PBS or SPB where indicated. The diet was replaced every 48 hours. The number of surviving aphids was assessed daily until one condition began producing nymphs. To determine if the treated survival curves were significantly different from the control the data was analysed by log-rank testing, the p-values are reported where relevant. Based on the survival percentages from day five, the LC_{50} value was calculated, by nonlinear regression (Graphpad Prism 8), for each protein against each insect species. The LC_{50} of a compound is the concentration which results in 50 % mortality of the test insects during the assay period. All feeding assays for GNA in this section were performed by Dr Prashant Pyati. All of the survival assays were performed with samples derived from the pH 4.5 fermentations of both PAF/GNA and PAF (Chapter 5).

6.2.1 Oral insecticidal activity of PAF, GNA and PAF/GNA on *Acrythosiphon pisum*

When *A. pisum* were fed on recombinant proteins, a dose-dependent response in survival was observed in all experiments, i.e. as concentration decreased aphid survival increased. As shown in Figure 6.1A, control survival was 92 % whereas 100 % mortality was recorded for aphids fed PAF at concentrations of 1.0 mg mL^{-1} and 0.75 mg mL^{-1} on day six and nine, respectively. No other protein condition resulted in complete mortality before the end of the assay, but survival decreased relative to the control. All treated curves were significantly different from the control survival in all conditions: 1.0 – 0.25 mg mL^{-1} ($p < 0.0001$), 0.1 mg mL^{-1} ($p = 0.023$). When GNA was fed at concentrations between 0.8 and 0.1 mg mL^{-1} complete mortality occurred at 0.8 mg mL^{-1} (day nine) and for 0.6 mg mL^{-1} (day 12) treatments as compared to control survival of 92 %. GNA concentrations of < 0.4 mg mL^{-1} caused a reduction in survival but not complete mortality. All survival curves were significantly different from the control: 0.8 – 0.2 mg mL^{-1} ($p < 0.0001$), 0.1 mg mL^{-1} ($p = 0.0008$). For PAF/GNA assays preliminary tests indicated that PAF/GNA had to be fed at lower concentrations as compared to PAF or GNA to allow LC_{50} values to be calculated. Day one nymphs were fed PAF/GNA at final dietary concentrations of: 0.4, 0.3, 0.2, 0.1 or 0.04 mg mL^{-1}. As seen in Figure 6.1C, control survival was 100 %. The high concentrations of PAF/GNA, 0.4, 0.3 and 0.2 mg mL^{-1} resulted in 100 % mortality on days three, six and ten, respectively. The remaining conditions resulted in decreased survival

compared to the control but did not cause 100 % mortality. The treated survival curves were compared to the control, and all conditions were found to be significantly different: 0.4 – 0.1 mg mL^{-1} ($p < 0.0001$), 0.04 mg mL^{-1} ($p = 0.017$).

6.2.2 Oral insecticidal activity of PAF GNA and PAF/GNA on *Myzus persicae*

In all *M. persicae* feeding assays, a dose-dependent response was visible for all proteins. Recombinant PAF was fed to *M. persicae* at final dietary concentrations of 2.0, 1.5, 1.0, 0.5 and 0.1 mg mL^{-1}. The survival of the aphids over time is shown in Figure 6.2A. When fed at a concentration of 2.0 and 1.5 mg mL^{-1} survival dropped to 0 % by day eight. Treatments fed at <1.5 mg mL^{-1} did not result in complete mortality but showed reduced survival compared to the control. All PAF treatments were significantly different to the control: 0.25 mg mL^{-1} ($p = 0.026$), 0.5 mg mL^{-1} to 1.0 mg mL^{-1} ($p < 0.0001$). Figure 6.2B shows the survival data for GNA fed at between 0.1 mg mL^{-1} and 0.8 mg mL^{-1}. In this assay control survival was 96% whereas the survival of aphids fed on GNA at 0.8, 0.6, 0.4, 0.2 and 0.1 mg mL^{-1} was 20.0 %, 44.0 %, 72.0 % and 88.0 %, respectively. Compared to the control survival all GNA treatments were significantly different: GNA 0.2 mg mL^{-1} ($p = 0.0396$), GNA 0.4 mg mL^{-1} ($p = 0.0044$), GNA 0.6 and 0.8 mg mL^{-1} ($p < 0.0001$). Figure 6.2C shows the results of feeding PAF/GNA at final dietary concentrations of between 0.1 mg mL^{-1} to 1.0 mg mL^{-1} . Whilst control survival was 100 % on day 11 complete mortality was recorded in the PAF/GNA 1.0 and 0.75 mg mL^{-1} treatments by day eight and ten, respectively. Dietary concentrations of < 0.75 mg mL^{-1} caused a reduction in survival did not result in complete mortality. The PAF/GNA treatments 0.25 mg mL^{-1} ($p = 0.029$) and 1.0 to 0.5 mg mL^{-1} curves ($p < 0.0001$) were significantly different to the control.

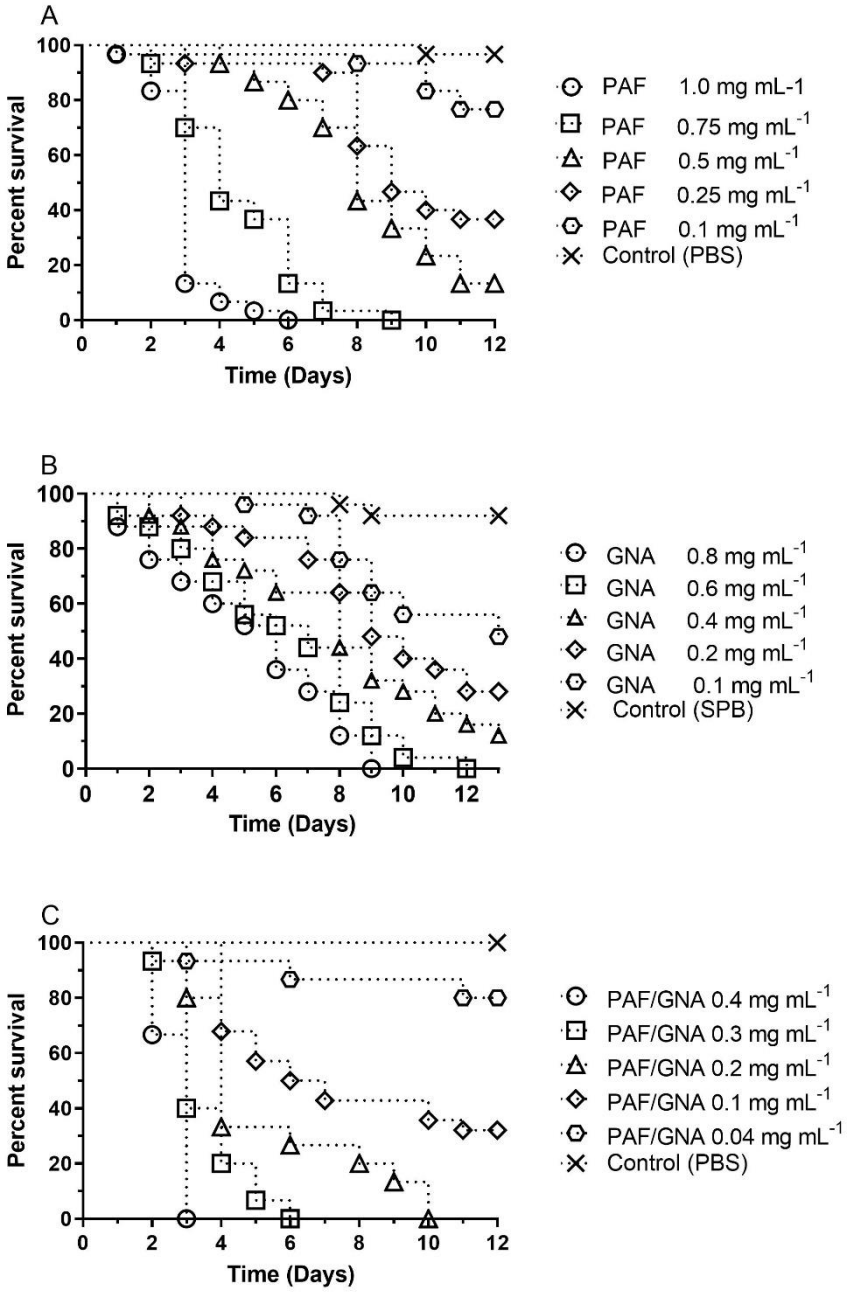


Figure 6.1 Survival of *A. pisum* fed on recombinant PAF, GNA or PAF/GNA. Purified recombinant proteins re-suspended in PBS were incorporated into artificial diet to give final concentrations as shown. The control treatment consisted of the diet with an equivalent volume of added PBS. A: Day one nymphs were fed on artificial diet containing 1.0, 0.75, 0.5, 0.25 and 0.1 mg mL^{-1} PAF. B: Day one nymphs were fed on artificial diet containing 0.8, 0.6, 0.4, 0.2, 0.1 mg mL^{-1} GNA. C: Day one nymphs were fed on artificial diet containing 0.4, 0.3, 0.2, 0.1, 0.04 mg mL^{-1} PAF/GNA. Each treatment consisted of 2 replicates of 15 nymphs. The GNA survival curve was generated from data produced by Dr Prashant Pyati.

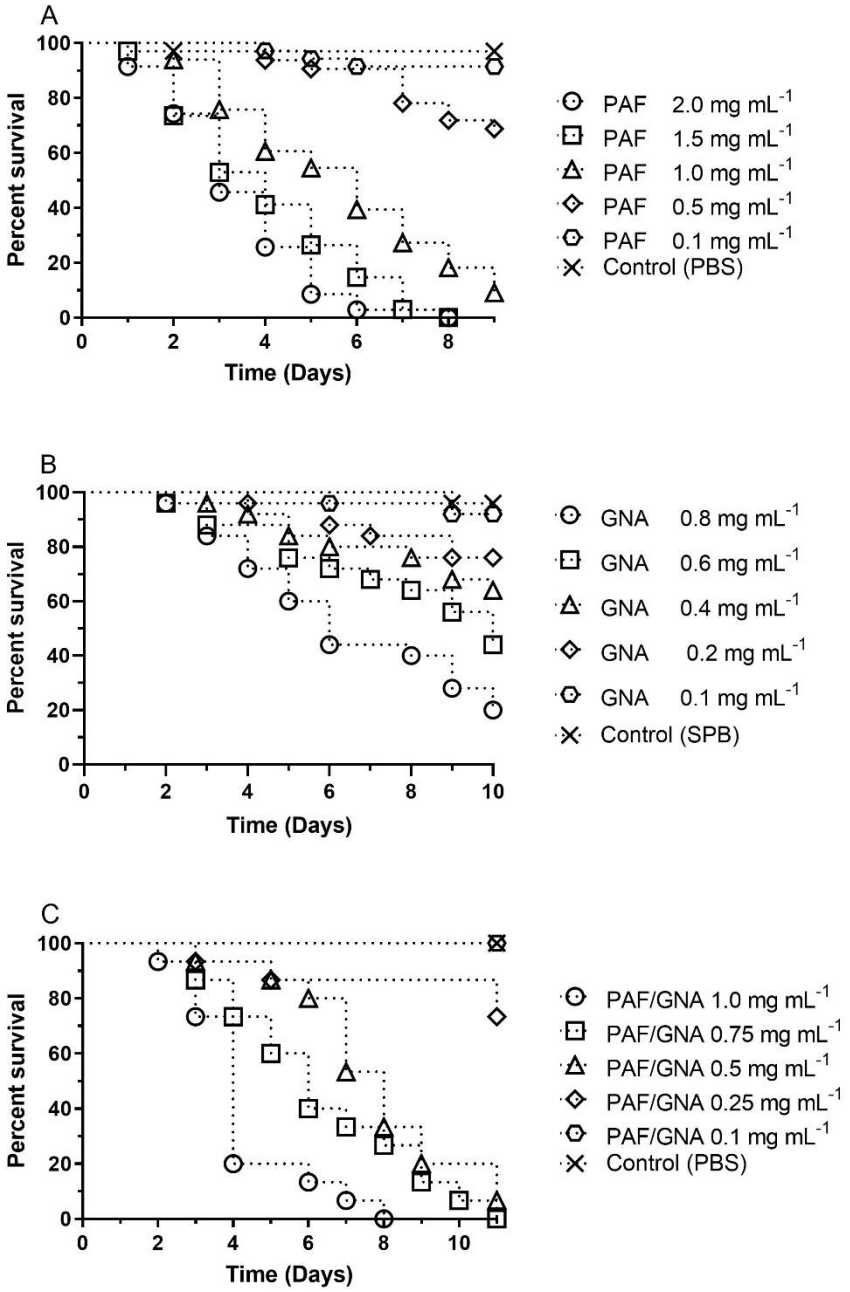


Figure 6.2 Survival of *M. persicae* fed on recombinant PAF, GNA or PAF/GNA. Purified recombinant proteins re-suspended in PBS were incorporated into artificial diet, to give the final concentrations shown. The control treatment consisted of the diet with an equivalent volume of PBS added. A: Day one and two nymphs were fed on artificial diet containing 2.0, 1.5, 1.0, 0.5, and 0.1 mg mL^{-1} PAF. B: Day one nymphs were fed on artificial diet containing 0.8, 0.6, 0.4, 0.2, 0.1 mg mL^{-1} GNA. C: Day one and two nymphs were fed on artificial diet containing 1.0, 0.75, 0.5, 0.25 and 0.1 mg mL^{-1} PAF/GNA. Each treatment consisted of 2 replicates of 15 nymphs. The GNA survival curve (B) was generated from data produced by Dr Prashant Pyati.

6.2.3 Oral insecticidal activity of PAF, GNA and PAF/GNA on *Sitobion avenae*

Recombinant PAF or PAF/GNA were fed to one to two-day old nymphs of *S. avenae,* at final dietary concentrations of: 1.0, 0.75, 0.5, 0.25, 0.1 mg mL^{-1}. GNA was fed at 0.8, 0.4, 0.2, 0.1 mg mL^{-1}. As shown in Figure 6.3, a dose-dependent response was visible in all three assays. The control survival was > 90 % in all assays. Complete mortality was recorded for PAF at concentrations of 1.0, 0.75 and 0.5 mg mL^{-1} on day four, six and seven, respectively. All PAF survival curves were significantly different to the control, PAF 0.1 mg mL^{-1} ($p = 0.0138$), 0.25 mg mL^{-1} ($p = 0.0019$), 0.5 mg mL^{-1}, 0.75 mg mL^{-1} and 1.0 mg mL^{-1} ($p < 0.0001$). Whilst none of the GNA concentrations tested resulted in complete mortality, survival, relative to controls was reduced by 30.0 %, 20.0 %, 10.0 % and 5.0 % after 7 days of feeding on GNA at 0.8, 0.4, 0.2 and 0.1 mg mL^{-1}. Only GNA at 0.4 mg mL^{-1} ($p = 0.013$) and 0.8 mg mL^{-1} ($p = 0.0003$) were found to be significantly different from the control condition. By contrast, PAF/GNA at concentrations of 1.0, 0.75 and 0.5 mg mL^{-1} resulted in 100 % mortality by days four, five and six, respectively. Dietary concentrations of < 0.5 mg PAF/GNA mL^{-1} caused a reduction in survival relative to the control treatment but did not cause total mortality. Survival curves for PAF/GNA at 1.0, 0.75 and 0.5 mg mL^{-1} were significantly different to control survival ($p < 0.0001$).

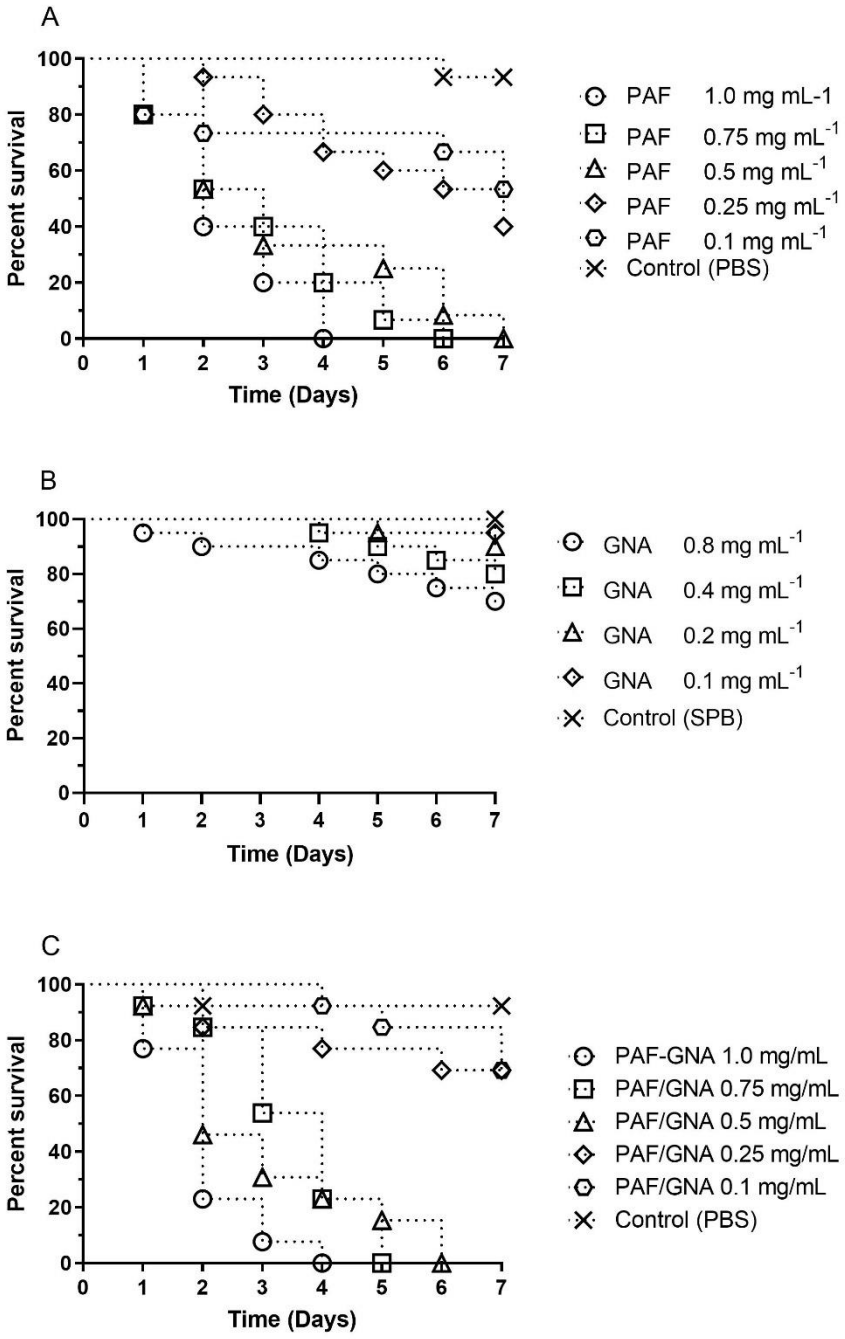


Figure 6.3 Survival of *S. avenae* fed on recombinant PAF, GNA or PAF/GNA. Purified recombinant proteins re-suspended in PBS were incorporated into the artificial diet to give the final concentrations shown, the control consisted of the diet with an equivalent volume of PBS added. A: Day 1 nymph aphids were fed on artificial diet containing 2.0, 1.5, 1.0, 0.5, and 0.1 mg mL^{-1} PAF. B: Day 1 nymphs were fed on artificial diet containing 0.8, 0.4, 0.2, 0.1 mg mL^{-1} GNA. C: Day 1 nymph aphids were fed on artificial diet containing 1.0, 0.75, 0.5, 0.25 and 0.1 mg mL^{-1} PAF/GNA. Each treatment consisted of 2 replicates of 13 nymphs. The GNA survival curve (B) was generated from data produced by Dr Prashant Pyati.

6.2.4 LC_{50} Comparison

The LC_{50} values were calculated from survival percentages on day five, by nonlinear regression using prism 8 as described in section 2.15. The software, (Graphpad Prism 8) fits the log of the survival data to a curve, and in instances where survival did not drop below 50 % extrapolates the result, this reduces the accuracy of the value as seen in the large confidence interval of the GNA condition of *S. avenae,* this is obviously not ideal, but in this instance it is the best available. These values and the 95 % confidence intervals (CI) were converted from mg mL^{-1} to nmol mL^{-1} to enable a direct comparison of protein toxicity on a molar basis rather than based on the mass of protein. Overlapping 95 % CIs were taken to mean a lack of significant difference between LC_{50}s. The LC_{50}s and 95 % CI values has been collated into Table 6.1. When fed to *A. pisum,* PAF/GNA had a significantly lower LC_{50} than either of the two constituent proteins (PAF and GNA) 12 and 12.9 times lower, respectively. In *M. persicae* PAF/GNA had an LC_{50} 3.5 times lower than PAF and 2.3 times lower than GNA. As the 95 % CIs from the LC_{50} calculations did not overlap LC_{50} values for PAF and PAF/GNA were deemed to be significantly different for both *A. pisum* and *M. persicae*. *Sitobion avenae* was the exception to this as the 95 % CIs overlapped for PAF/GNA (LC_{50} 15 nmol mL^{-1}) and PAF (LC_{50} 36 nmol mL^{-1}). The LC_{50} for GNA against *S. avenae* was found to be substantially higher than the LC_{50} of PAF/GNA or PAF. Recombinant PAF was found to be most active against *S. avenae* (LC_{50} 36 nmol mL^{-1}), followed by *A. pisum* (LC_{50} 58 nmol mL^{-1}) and *M. persicae* (115 nmol mL^{-1}). Recombinant GNA was observed to be similarly toxic to both *A. pisum* (LC_{50} 62 nmol mL^{-1}), *M. persicae* (LC_{50} 78.6 nmol mL^{-1}), but substantially less effective when fed to *S. avenae* (LC_{50} 186 nmol mL^{-1}). Finally, recombinant PAF/GNA was most effective against *A. pisum* (LC_{50} 5 nmol mL^{-1}) followed by *S. avenae* (LC_{50} 15.0 nmol mL^{-1}) and least effective against *M. persicae* (LC_{50} 33 nmol mL^{-1}). The LC_{50}s for PAF and PAF/GNA were significantly different in all species, whereas only the *S. avenae* GNA LC_{50} differed significantly from the other species (Table 6.1).

Table 6.1 LC_{50} values (day five) and 95 % confidence intervals for PAF, GNA and PAF/GNA against *A. pisum, M. persicae*, and *S. avenae*. GNA data was calculated based on feeding assays performed by Dr Prashant Pyati.

Protein	Pea aphid (*A. pisum)* nmol mL^{-1}		Peach-potato aphid (*M. persicae)* nmol mL^{-1}		Grain aphid (*S. avenae)* nmol mL^{-1}	
	LC_{50}	95% CI	LC_{50}	95% CI	LC_{50}	95% CI
PAF	58	52 - 94	112	111 – 114	22	10 – 22
GNA	62	50 - 77	79	65 – 94	181	102 - 755
PAF/GNA	5	3 - 7	34	29 - 38	15	13 - 36

6.3 Effects on fecundity of aphids fed on recombinant pea albumin (PAF) or PAF/GNA

The effect of PAF/GNA and PAF on the fecundity of *A. pisum* and *M. persicae* was assessed by counting nymphs produced by aphids fed on recombinant proteins up to day 16 after the onset of the assay; total nymphs were divided by the number of non-winged adults present at the beginning of nymph production in the survival assays presented previously. Results presented in Figure 6.4, show that adults fed on PAF or PAF/GNA produce lower numbers of nymphs at the higher sublethal concentrations tested. All feeding regimes showed a dose-response, with decreasing nymph production as protein concentration increased. Feeding PAF or PAF/GNA to *A. pisum* or *M. persicae* caused a 100 % reduction in fecundity relative to the control at the highest concentration tested in each experiment. Conversely, the lowest concentration of PAF fed to *A. pisum* appeared to increase fecundity slightly over the control. PAF/GNA appeared to have a greater effect on *A. pisum* and *M. persicae* fecundity than PAF as when fed at the same concentration of 0.1 mg mL^{-1} the PAF treatment produced a mean of 2.7 nymphs per adult, whilst PAF/GNA produced no nymphs. Similarly, when fed to *M. persicae* PAF and PAF/GNA 0.1 mg mL^{-1} treatments each produced a mean of 3.8 and 3.1 nymphs per adult, respectively.

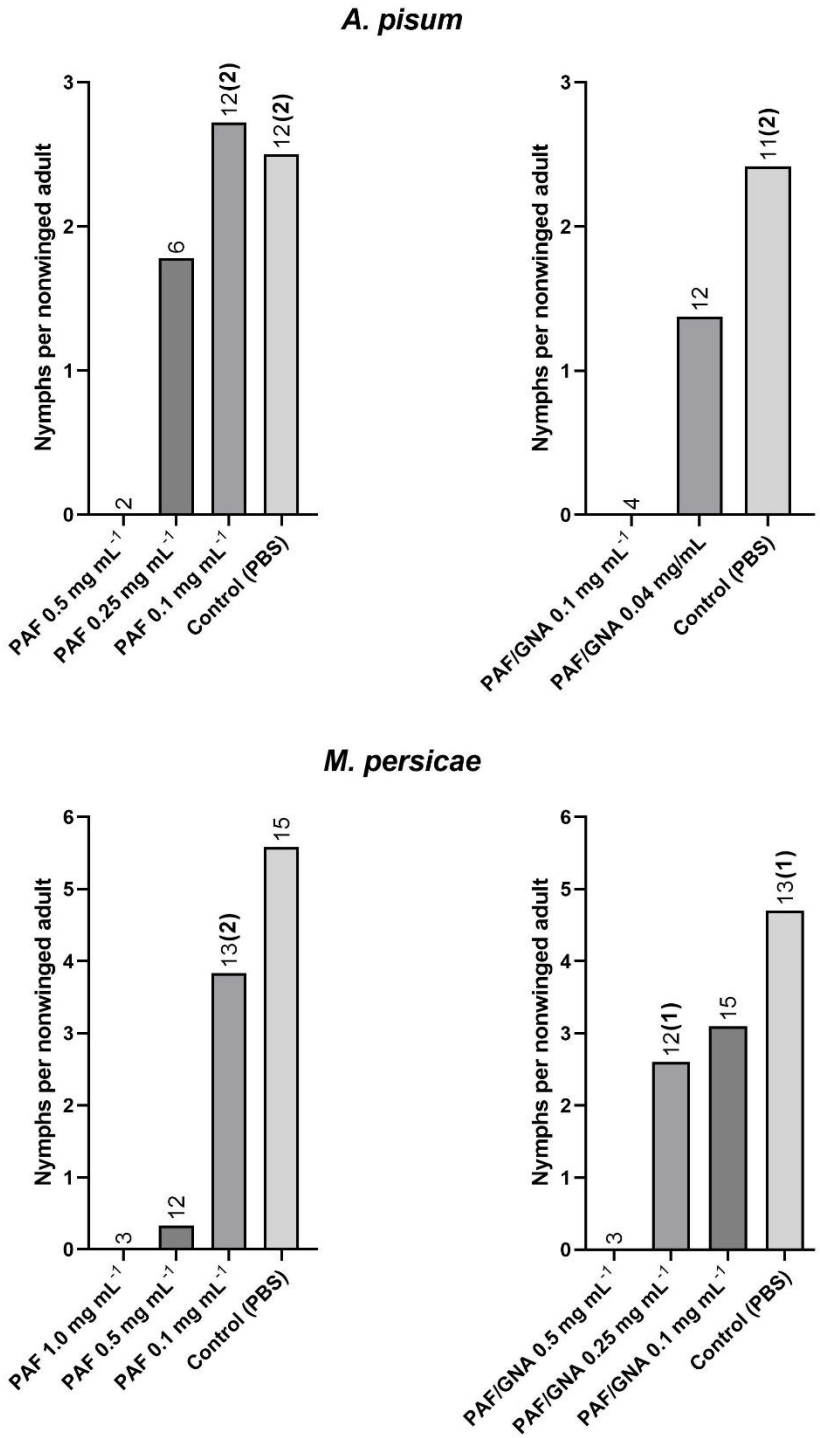


Figure 6.4 The effect of PAF and PAF/GNA on the fecundity of *A. pisum* and *M. persicae*. The bars represent the number of nymphs produced per non-winged adult by day 16 of the feeding assays. The number stated above the bar shows the number of non-winged (winged) adults present at the beginning of nymph production. This data was produced from a single feeding assay for each protein-species combination.

6.4 Retention of PAF and PAF/GNA in the *A. pisum* gut

As demonstrated by the LC_{50}s, PAF/GNA is 12 times more active against *A. pisum* than PAF alone. GNA could be enhancing activity by anchoring PAF in the vicinity of V-ATPase proton pumps in the gut epithelium, through binding to mannose containing glycoproteins present on the lumen surface. To investigate if differences existed in the persistence and/or binding of the different proteins to the aphid gut epithelia a pulse-chase experiment was conducted. To this end, PAF, PAF/GNA, GNA and ovalbumin were labelled with fluorescein isothiocyanate (FITC), with ovalbumin acting as a control as it should not interact with the epithelial membrane. Five-day old *A. pisum* were fed on diets containing FITC labelled proteins (8 nmol mL^{-1} of PAF, 8 nmol mL^{-1} PAF/GNA, 8 nmol mL^{-1} of GNA or 5 nmol mL^{-1} of ovalbumin [Sigma-Aldrich, St. Louis, USA]) and propidium iodide (100 ng μL^{-1}), as a counterstain, for 24 hours. They were subsequently fed on artificial diet containing propidium iodide (100 ng μL^{-1}), but without fluorescently labelled proteins, for 24 or 48 hours. The guts were dissected and imaged immediately. PAF samples used in this experiment were derived from the pH 5.0 fermentation, PAF/GNA from the pH 4.5 glucose fermentation and GNA from the glycerol fermentation (Chapter 5, Chapter 3).

In all conditions, where visible, the protein appeared to be localised to the bulbous structure of the anterior midgut. Representative images for each condition are shown in Figure 6.5. The control and ovalbumin fed conditions showed no visible fluorescence; and only minor fluorescence was detected when the pixel intensity was analysed using FIJI (Figure 6.6). Fluorescence is visible in both PAF and GNA fed conditions after chase feeding for 24 hours, and reduced fluorescence is visible in guts extracted after 48 hours. The intensity of the fluorescence was much higher in the PAF/GNA condition after both 24 and 48-hour chase feeding. This pattern is also observed when the mean pixel intensity is quantified across multiple dissected samples (Figure 6.6). Mean pixel intensity was quantified as described in section 2.24. The mean intensity of the traced area was then averaged, unpaired Welch's t-tests were performed to determine significant difference compared to the PAF/GNA condition. PAF/GNA was found to be significantly more intense than all the other conditions tested across both time points (24 hours: $p < 0.0005$; 48 hours: $p < 0.05$; Welch's t-test).

In initial experiments PAF and GNA displayed substantial fluorescence when imaged after 24 and 48 hours. However due to the increased fluorescence observed in the PAF/GNA treatments, laser power had to be reduced to prevent saturation of the PAF/GNA images and consistency across conditions was required to allow accurate measurement of pixel intensity across all images using standardised imaging settings.

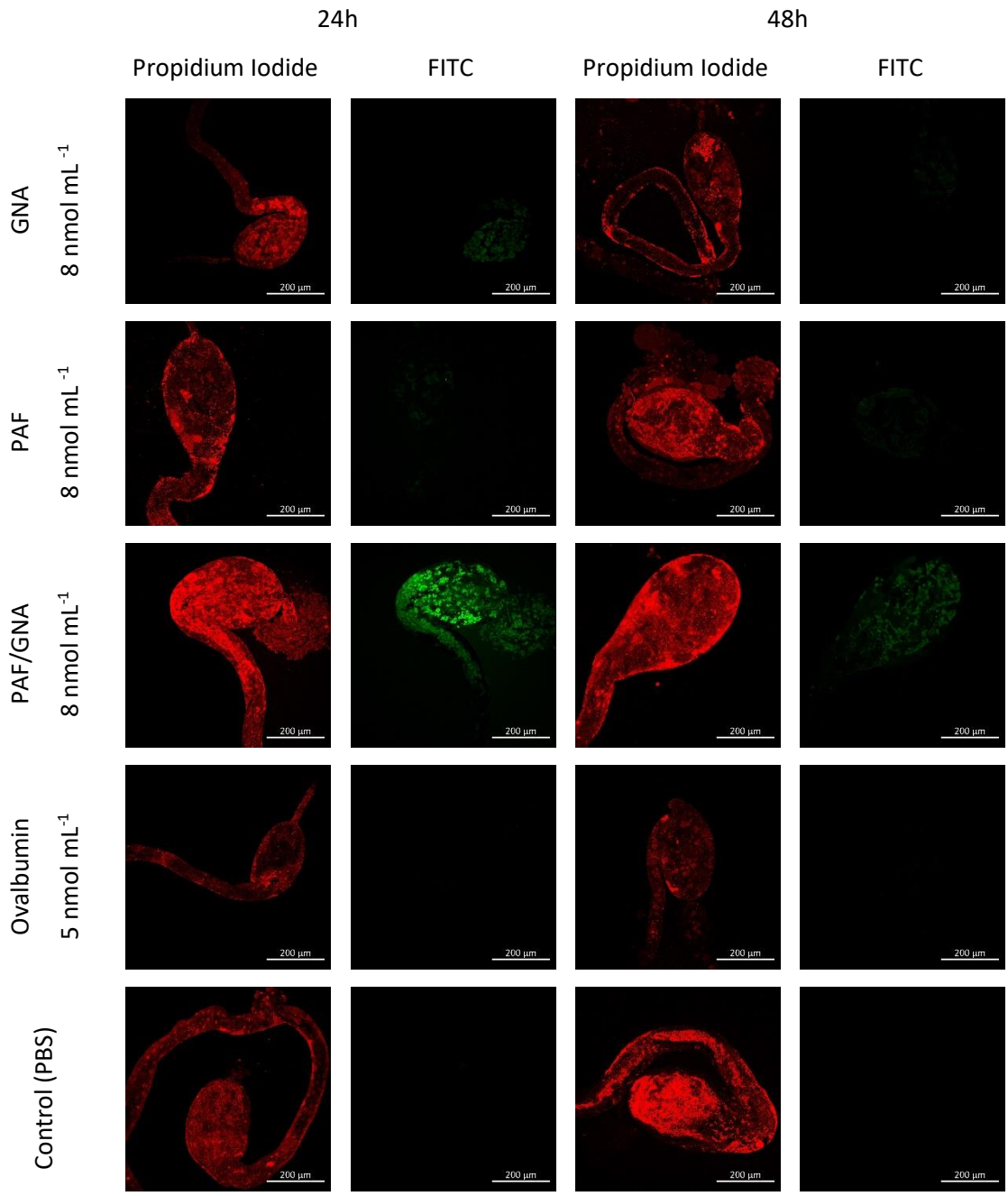


Figure 6.5 Confocal images of *A. pisum* dissected guts in a pulse-chase feeding experiment with FITC labelled proteins. Day five aphids were fed for 24 hours on an artificial diet containing FITC labelled protein and propidium iodide (counterstain). Aphids were then chase fed with artificial diet plus propidium iodide for 24 or 48 hours. Guts were dissected in sterile ice-cold PBS and imaged immediately using a Zeiss 880 LSM. The following settings were used to capture all images above: Objective: Plan-Apochormat 20x/0.8 M27, propidium iodide was excited with a 543 nm laser on 1.5% power, wavelengths were collected between 584-735 nm, gain: 800. FITC was excited with a 488 nm on 0.5% power; wavelengths were collected between 493-572nm, gain: 575.

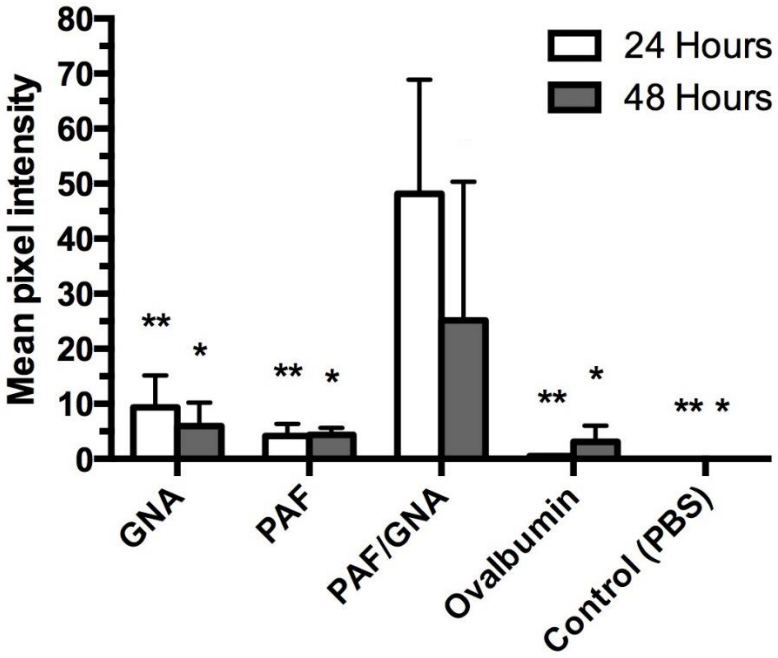


Figure 6.6 Average pixel intensity of *A. pisum* gut from confocal imaging of a pulse-chase experiment with FITC labelled proteins. GNA, PAF and PAF/GNA were fed at 8 nmol mL^{-1} and ovalbumin was fed at 5 nmol mL^{-1}. Pixel intensity was measured using the measure function of FIJI, the gut was traced using the outline provided by the propidium iodide counterstain and then fluorescence in the green wavelength was measured. The mean pixel intensity was taken for a minimum of six dissected guts per treatment and standard error was calculated. Pairwise T-testing was used to determine statistical significance * = $p < 0.05$ when compared to PAF/GNA. ** = $p < 0.0005$. Error bars = standard error.

6.5 Stability of PAF, GNA and PAF/GNA when exposed to *A. pisum* and *M. persicae* salivary secretions in the diet

The stability of the recombinant proteins within the diet mixture was assessed when exposed to salivary secretions from *A. pisum* and *M. persicae*. This was done by collecting diet spiked with recombinant protein at four time points over 48 hours (4 h, 8 h, 24 h and 48 h) of continuous aphid feeding and western blotting. A no-aphid control was also maintained under the same conditions. Recombinant proteins used in this assay were derived from the same fermentations as in section 6.4.

Recombinant PAF is moderately stable in the presence of aphid salivary secretions. In Figure 6.7, immunoreactive bands at the predicted mass were visible after 48 hours; no other immunoreactive bands were visible. After 48 hours of exposure to both *A. pisum* and *M. persicae* salivary secretions PAF showed reduced intensity by approximately 50 %. This is likely due to the cleavage of the C-terminal his-tag at the assay temperature (20 °C) as the no aphid 48-hour control also has reduced immunoreactivity. This is not unexpected as the his-tag on the C-terminus is likely exposed and therefore prone to cleavage by environmental proteases.

The stability of recombinant GNA exposed to the salivary secretions of day five *A. pisum* and *M. persicae* is shown in Figure 6.8. An immunoreactive band for both anti-GNA and anti-His antibodies can be seen at 14 kDa, the predicted molecular weight of GNA. After 48 hours of feeding the intensity of the immunoreactive band was approximately equal to that of the 0 hours no aphid control. There is a decrease in intensity after 48 hours in Figure 6.8B, which is likely due to proteolytic cleavage of the his-tag. A dark immunoreactive strip across the blot at approximately 20 kDa in Figure 6.8A, is suspected to be GNA contamination of one of the components involved in running sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels.

The immunointensity of PAF/GNA was stable over time in diet samples exposed to *A. pisum* or *M. persicae* salivary secretions, as displayed in Figure 6.9A. There was a slight decrease in immunointensity at the 48-hour time point in *A. pisum* that was not observed in *M. persicae*. *M. persicae* showed a decrease in intensity between 4 and 8 hours then remained stable. By contrast, when probed with anti-his antibodies PAF/GNA was not detectable after 4 hours of *M. persicae* feeding, suggesting that salivary secretions contain proteases capable of cleaving the his-tag present at the C-terminus of PAF/GNA.

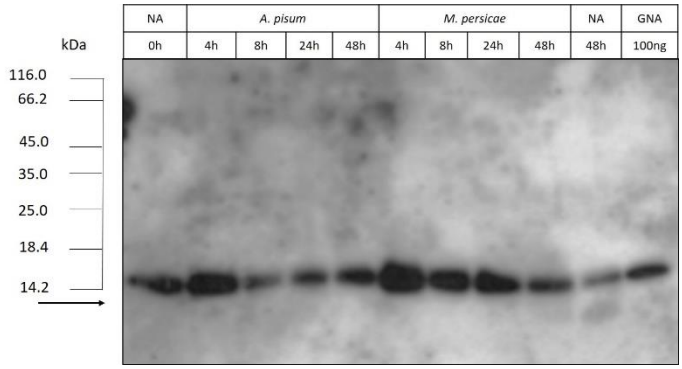


Figure 6.7 Stability of PAF to degradation by aphid salivary secretions over time. Western blot probed with anti-His antibodies of artificial diet samples containing PAF after *A. pisum* and *M. persicae* feeding for designated time points (n=15 aphids per treatment). NA represents diet plus PAF control (no aphid feeding) and recombinant GNA was used as a positive control. Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

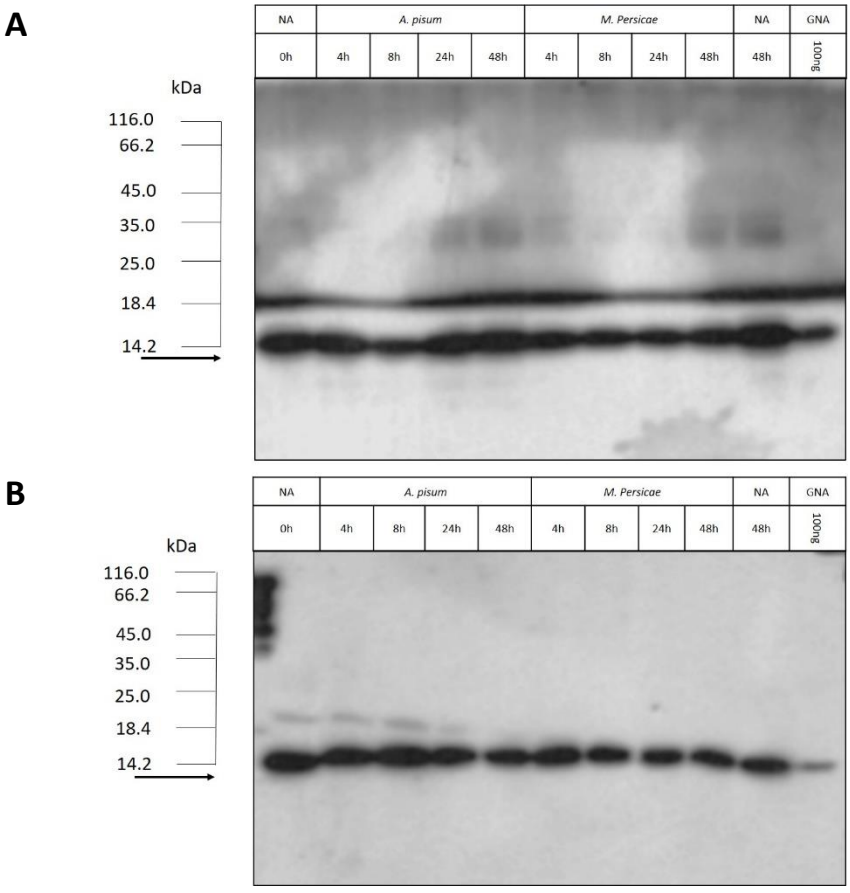


Figure 6.8 Stability of GNA to degradation by aphid salivary secretions over time. Western blot of artificial diet samples containing GNA after *A. pisum* and *M. persicae* feeding for designated time points probed with anti-GNA (A) and anti-His (B) antibodies. NA represents diet plus GNA control (no aphid feeding). Recombinant GNA was used as a positive control. Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

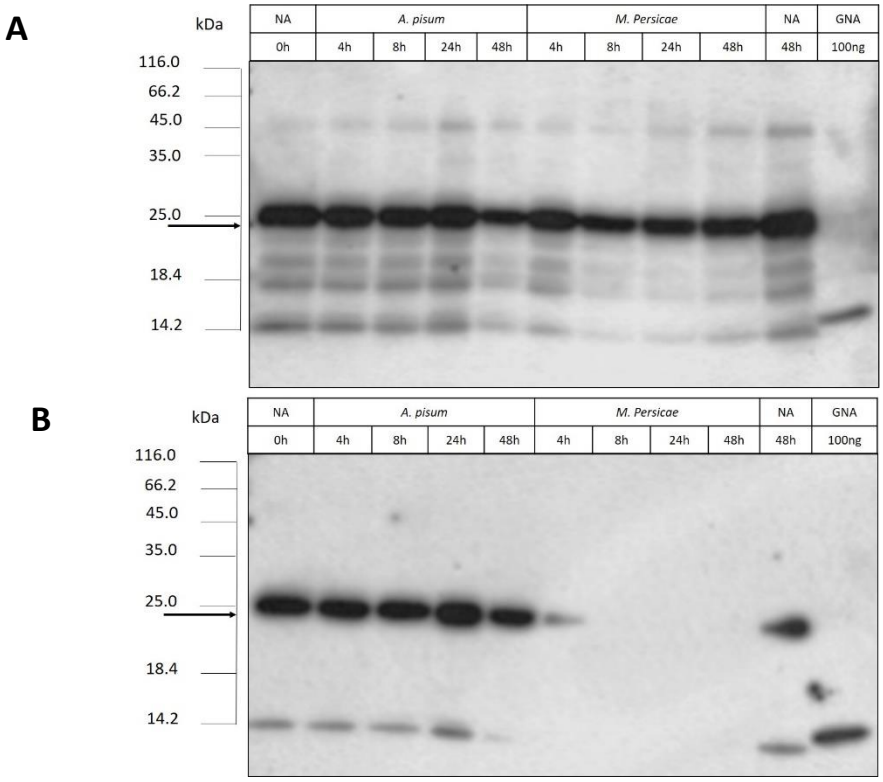


Figure 6.9 Stability of PAF/GNA to degradation by aphid salivary secretions over time. Western blot of artificial diet samples containing PAF/GNA after *A. pisum* and *M. persicae* feeding for designated time points. Blots probed with anti-GNA (A) and anti-His (B) antibodies. NA represents diet plus PAF/GNA control (no aphid feeding). Recombinant GNA was used as a positive control. Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

6.6 *In vivo* stability of orally delivered PAF, GNA and PAF/GNA in *A. pisum* and *M. persicae*

The stability of PAF, GNA and PAF/GNA *in vivo* was assessed by continuously feeding proteins (at final dietary concentrations of 0.4 mg mL^{-1}) in artificial diet to five day old *A. pisum* and *M. persicae*. Fifteen aphids per time point for each treatment were flash frozen at each 4, 8, 24 and 48 hours. Protein stability *in vivo* was analysed by western blotting of 60 µg total aphid protein extract, estimated by BCA assay. Recombinant proteins used in this assay were derived from the same fermentations as in section 6.4. Assessment of the stability of PAF *in vivo* was not possible as no protein was detectable in any protein extract when probed with anti-his antibodies, at any of the time points in this experiment, indicating that the histidine tag is rapidly cleaved by *A. pisum* and *M. persicae* gut proteases (Figure 6.10). GNA stability *in vivo* is shown in Figure 6.11. In Figure 6.11A, an immunoreactive protein was detected in aphid extracts with orally delivered GNA at approximately 13 kDa. This is approximately 1 kDa smaller than the predicted mass of GNA, suggesting some cleavage of the his-tag by aphid gut proteases in both species. Generally, the intensity of the immunoreactive band increased over time in both species (indicative of retention of ingested protein), although it increased at a greater rate in *M. persicae* treatments up to 24 hours. At 48 hours the *M. persicae* extract shows decreased intensity of the immunoreactive band corresponding to PAF/GNA cleavage products. The decrease is possibly related to reductions in feeding. A cross reactive protein was visible at 45 kDa in all lanes including time 0. The presence of an immunoreactive strip across the blot at approximately. The 18 kDa protein is suspected to be GNA contamination from one of the components involved in running SDS-PAGE gels. As with PAF, GNA was not detectable by probing with anti-his antibodies providing further evidence that the his-tag is highly susceptible to proteolytic cleavage (Figure 6.11B).

Intact fusion protein was not visible at any time point in Figure 6.12A. After 4 hours of *A. pisum* feeding immunoreactive bands are visible at 20 kDa and approximately 13 kDa are presumed to be cleavage products of PAF/GNA. From 8 hours onward the 20 kDa band is not detected and the 13 kDa band fades. This is the opposite to observations in *M. persicae*, two immunoreactive bands at approximately 13 and 14 kDa are present in all time points. Unfortunately, due to an artefact 4 h and 8 h timepoints were partially obscured, but intensity of the low MW band appears to increase with time, whilst the 14 kDa band remains stable. A cross-reactive band between 45-60 kDa was observed in *M. persicae* samples, this is expected in *A. pisum* timepoints, however incomplete protein transfer occurred at high molecular weights in these lanes, based on the background signal. A black line at approximately 18 kDa is visible in all lanes' this is present due to GNA contamination of one of the SDS-PAGE components.

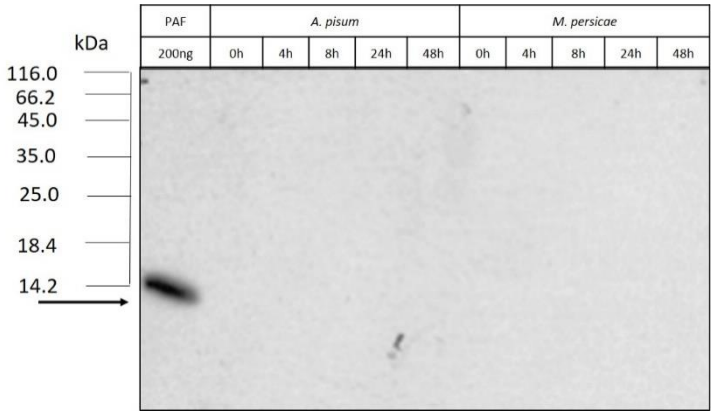
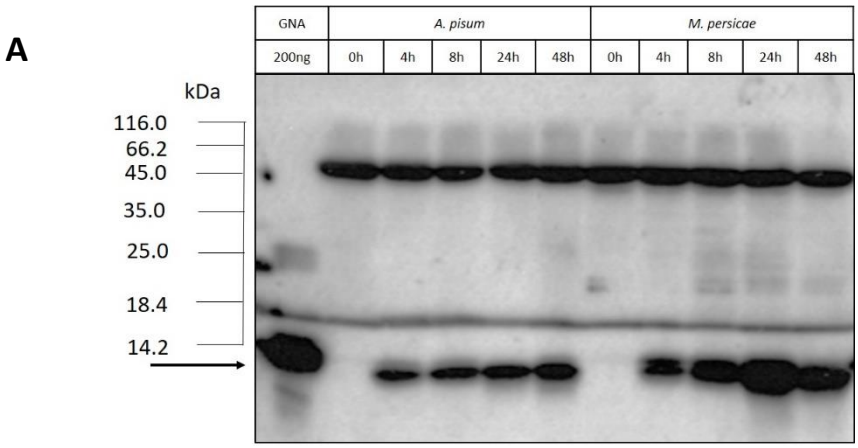


Figure 6.10 *In vivo* stability of PAF to proteolysis in the gut of *A. pisum* and *M. persicae*. Western blot probed with anti-His antibodies of total aphid protein extracts (60 µg per lane) after feeding on artificial diet containing 0.4 mg mL^{-1} PAF for designated time points. *2*00 µg PAF was used as a positive control. Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.



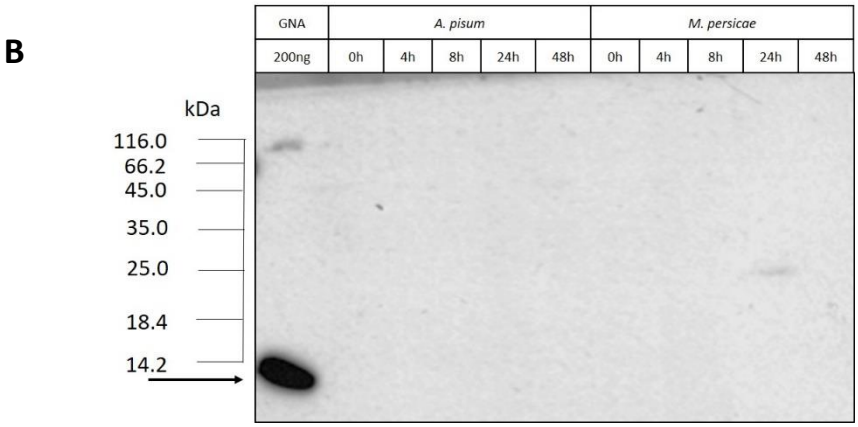


Figure 6.11 *In vivo* stability of GNA to proteolysis in the gut of *A. pisum* and *M. persicae*. Western blot probed with anti-GNA (A) and anti-His (B) antibodies of total aphid protein extracts (60 µg per lane) after feeding on artificial diet containing 0.4 mg mL^{-1} GNA for designated time points 200 ng GNA was used as a positive control. Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

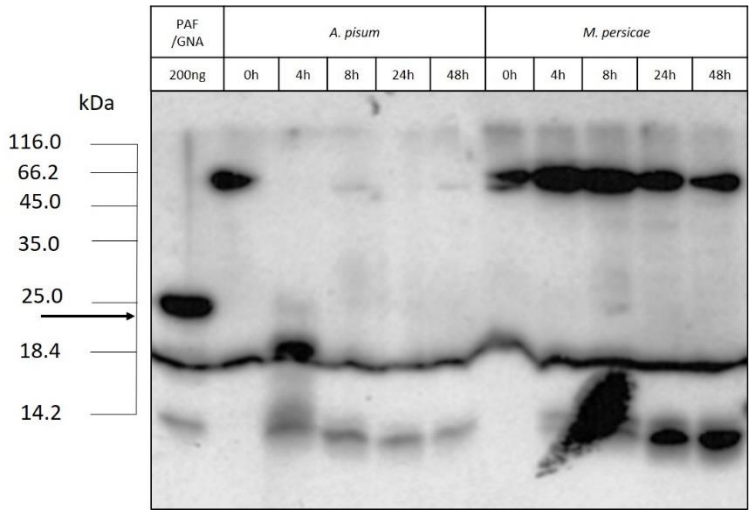


Figure 6.12 ***In vivo*** **stability of PAF/GNA to proteolysis in the gut of** ***A. pisum*** **and** ***M. persicae*****.** Western blot of aphid protein extracts (60 µg per lane) after feeding on artificial diet containing 0.4 mg mL^{-1} PAF/GNA for designated time points probed with anti-GNA antibodies. PAF/GNA was used as a positive control. Position of protein marker mix (Thermo) run on the same gel is depicted on the left-hand side. The arrow indicates the expected protein mass.

6.7 Free choice feeding assays

The target site of PA1b is a key component in nutrient transport and pH maintenance of the insect gut. Disruption of V-ATPase function might result in anti-feeding behaviour and cause aphids to seek new food sources. To investigate this day one *A. pisum* nymphs and day one and two *M. persicae* nymphs were supplied with a choice of two diets, one containing either PAF/GNA, PAF, GNA or ovalbumin (Sigma-Aldrich, St. Louis, USA) and a control containing PBS. Recombinant proteins used in this assay were derived from the same fermentations as in section 6.4. The two diets were placed on separate halves of a 3 cm diameter aphid feeding chamber sandwiched between two layers of parafilm (Figure 6.13B). The number of aphids that congregated on each diet after 24 and 48 hours were counted, and a mean calculated from six replicates of 20 aphids (Figure 6.13A). The mean number of *A. pisum* feeding on the control diet after 24 and 48 hours was significantly higher compared to all protein conditions ($p < 0.05$, Welch's t-test) except for ovalbumin where there was no significant difference in the number of aphids feeding on either diet ($p > 0.05$, with no added protein whether this be PAF, GNA, or PAF/GNA. *Myzus persicae* did not display a significant preference for any diet ($p > 0.05$, Welch's t-test), except after 48 hours there were more insects on control diet than on diet with PAF/GNA ($p = 0.0016$, Welch's t-test).

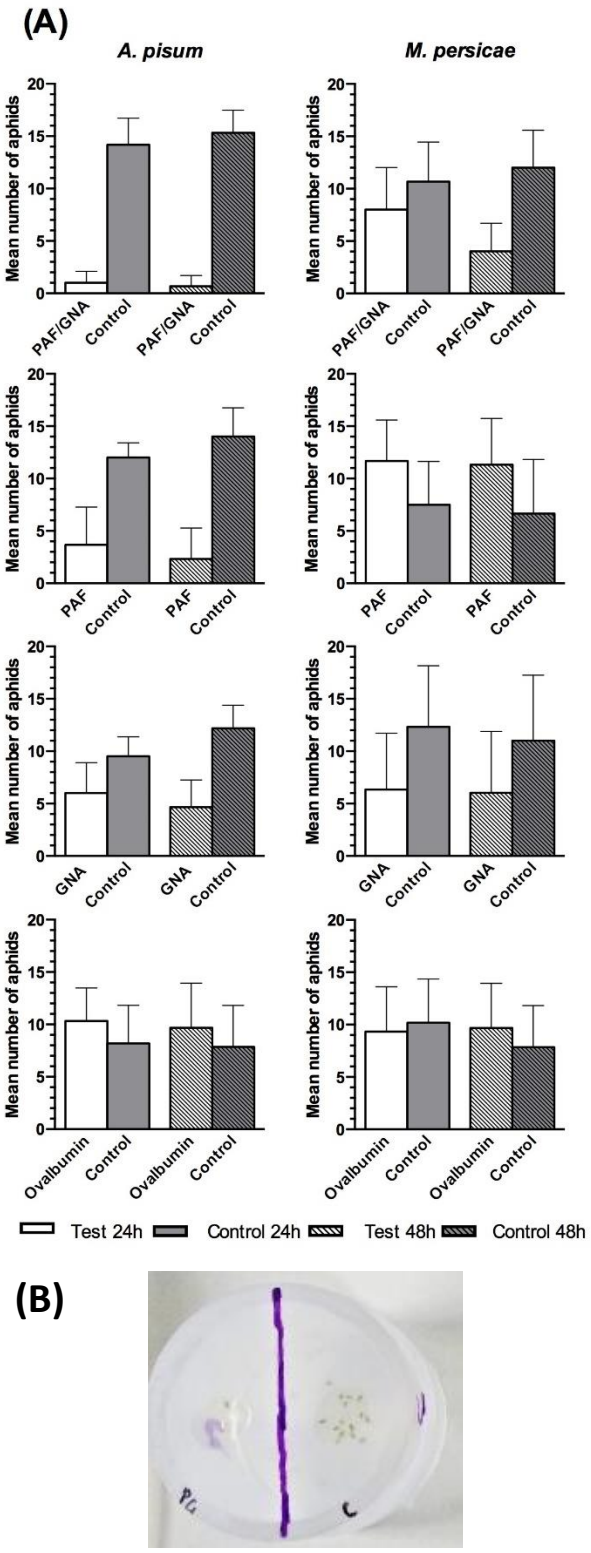


Figure 6.13 Mean number of aphids on each diet when given a choice between diet containing PAF/GNA, GNA and PAF. A: Shows the mean number of aphids present on each diet. PAF/GNA: 0.4 mg mL^{-1}, PAF: 0.2 mg mL^{-1} and GNA: 0.2 mg mL^{-1}. Bars represent the mean value of six replicates, n = 20. Error bars represent ± SD of the mean. B: An example of the free choice assay setup, protein containing diet are left of the line, control diet to the right, only aphids within the diet were counted as feeding on the diet.

6.8 Discussion

PA1b is an interesting protein for development as unlike traditional pesticides classes it does not target the insect nervous system but instead inhibits V-ATPase proton pumps, present in the insect gut epithelial membrane, which are responsible for nutrient uptake and ion balance (Beyenbach and Wieczorek, 2006; Chouabe *et al*., 2011; Wieczorek *et al*., 2000). It is thought to bind to the c-subunits present on the apical membrane surface and prevent rotation of the pump by binding to the static e-subunit, which is required for proton transport (Muench *et al*., 2014). Publications relating to the oral activity of PA1b in insects are limited to two papers and a patent one relating to weevils and one which assessed toxicity towards a number of different species (Delobel *et al*., 1998; Gressent *et al*., 2003, 2007). Pea flour was also found to be toxic to a large number of *Sitophilus* strains this has since been attributed to the presence of PA1b in the flour (Delobel *et al*., 1998). Further reference to expanded toxicity towards a range of species is referenced by personal communication in a review by Gressent *et al*., (2011).

This chapter provides evidence to demonstrate that linking PAF (PA1b+PA1a) to GNA enhances its biological activity when fed to the aphids *Acrythosiphon pisum* and *Myzus persicae* but not when fed to *Sitobion avenae*. *A. pisum* was significantly more susceptible than *M. persicae*. The effect on fecundity was also assessed, and dose-dependent effects on nymph production were observed. The increased retention of PAF/GNA compared to PAF or GNA in the *A. pisum* gut up to 48 hours after feeding was demonstrated using fluorescent confocal microscopy and provides an explanation for PAF/GNA increased toxicity. *In vivo* stability assays were carried out to determine if the reduced activity towards *M. persicae* is due to reduced protein stability when exposed to aphid salivary secretions or gut proteases. This was not found to be the case. Finally, antifeeding effects were assessed by conducting a free choice assay for each protein and assessing the number of aphids feeding on each protein after a period of 24 and 48 hours, *A. pisum* displayed a significant preference for the control diet over PAF/GNA, PAF and GNA.

Both recombinant PAF and PAF/GNA showed insecticidal activity when supplied in artificial diet to three different aphid species. Linking PAF to GNA led to a 12-fold decrease in LC_{50} against *A. pisum* and a 4-fold decrease of LC_{50} against *M. persicae* as compared to PAF alone, suggesting that the proteins are inhibiting V-ATPase proton pumps. Gressent *et al*., (2007) examined the oral insecticidal activity of PA1b, purified from pea seeds using reverse phase HPLC, against insects from various orders, including *A. pisum* and *M. persicae*. The molar concentration required by the authors to cause 100 % mortality in *A. pisum* (135 nmol mL^{-1}) after 6 days was somewhat higher than required in this study (PAF: 86.2 nmol mL^{-1}; PAF/GNA: 12.6

nmol mL^{-1}). The same study showed PA1b purified from pea seeds at 270 nmol mL^{-1} caused no mortality when fed to *M. persicae*, which is two-fold and eight-fold greater than the LC_{50} of PAF and PAF/GNA toward the same species. PA1b was also found to have an IC_{50} against Sf9 cells of 87 nmol mL^{-1} (Rahioui *et al*., 2014). The recombinant PAF (PA1b + PA1a) produced in this study is more effective against both *M. persicae* and *A. pisum* than PA1b purified directly from pea seeds. The additional protein in PAF, PA1a, is not toxic to Sf9 cells in its own regard and is unlikely to be directly contributing to toxicity (Chouabe *et al*., 2011; Higgins *et al*., 1986). One could hypothesize that this result is a consequence of its increased size, 11.6 kDa versus 3.7 kDa, PA1b functions by inhibiting proton transport by physically preventing the rotor movement by binding to the c-subunit and jamming against the e-subunit of V-ATPase (Muench *et al*., 2014). Whereas fusion of GNA to PAF resulted in significantly enhanced efficacy, as compared to PAF alone, towards *A. pisum* and *M. persicae* this was not the case for *S. avenae.* Speculation on this topic led to the hypothesis that perhaps *S. avenae* has very few mannose-glycoproteins in the gut epithelium. The reduced presence of mannosylated gut epithelial proteins would reduce the ability of GNA to bind to the gut epithelium and this may explain both the high LC_{50} of GNA and the lack of significant difference in insecticidal activity between PAF/GNA and PAF.

The calculated LC_{50} values from Dr Prashant Pyati's recombinant GNA feeding trails were 62 nmol mL^{-1}, 79 nmol mL^{-1} and 181 nmol mL^{-1} for *A. pisum, M. persicae* and *S avenae* respectively. These values are generally high compared to feeding trials of *A. pisum* and *M. persicae* performed with GNA purified from snowdrop (Rahbé *et al*., 1995b; Sauvion *et al*., 1996). This reduced toxicity may be due to the use of a C-terminally truncated version of GNA (105 residues) in the assays reported here versus the native (109 residue) protein, and this may interfere with the ability of GNA to form tetramers through C-terminal strand exchange and effect activity (Hester *et al*., 1995). These results suggest that the observed significantly greater efficacy of PAF/GNA as compared to GNA and PAF is not simply attributable to the additive effects of GNA and PAF but that there is a synergistic effect.

Controlling fecundity is a key factor to consider in aphid management as due to their parthenogenic lifestyle aphid populations can increase rapidly. *Myzus persicae* can begin reproducing 12 days after birth and one adult can produce up to 80 nymphs over its life time (van Emden *et al*., 1969). A single apterous *A. pisum* can produce upwards of 90 nymphs when feeding on *P. sativum* over the course of its 30 day lifetime with nymph production beginning between day 7 and 10 (Legrand and Barbosa, 2000). As compared to controls, dose-dependent decreases in fecundity were generally observed when *A. pisum* or *M. persicae* were fed on sub-lethal concentrations of PAF or PAF/GNA. Feeding on PAF/GNA resulted in greater reductions in

fecundity compared to feeding on PAF for both *A. pisum* and *M. persicae.* Despite the lack of literature on the effect of PA1b on fecundity, decreased nymph production is not surprising. RNAi mediated knockdowns of V-ATPase subunits A and E have resulted in significant reductions in the fecundity in another hemipteran pest *Cimex lectularius* (Basnet and Kamble, 2018). In general metabolic challenges to can reduce aphid fecundity; Douglas *et al*., (2006) demonstrated that both high and low concentrations of sucrose in artificial diets result in reduced fecundity and Watt (1979) observed that aphids feeding on senescent wheat leaves (reduced nutritional value) displayed reduced fecundity compared to aphids feeding on the ears or new leaves (greater nutritional value). Further, the fusion protein Hv1a/GNA and GNA both caused a decrease in *M. persicae* fecundity when fed at 0.25 mg mL^{-1} (Nakasu *et al*., 2014b). Further data from our lab demonstrates that sublethal concentrations of GNA and GNA based fusion proteins generally cause reduced fecundity in *M. persicae, 0.1*mg mL^{-1} concentrations of GNA, Pro-Hv1a/GNA, PI1a/GNA all caused a > 50 % reduction in nymphs per non-winged adult compared to the control (Pyati - unpublished data).

Localisation and retention of PAF/GNA at the gut epithelium was suspected to be the cause of the significant increase in toxicity observed when *A. pisum* was fed with PAF/GNA compared to PAF. Evidence for GNA localisation in the hemipteran gut has been shown previously (Powell *et al*., 1998; Yang *et al*., 2014). Confocal microscopy in conjunction with FITC labelled PAF/GNA, PAF, GNA and ovalbumin was used to examine the fluorescence retained in the gut in chase feed experiments. Retained fluorescence after clearance of diets containing proteins from the gut by chase feeding was assumed to be due to binding of labelled proteins to the gut epithelium. In preliminary experiments both PAF and GNA showed binding in the stomach (anterior midgut) as observed for PAF/GNA; however imaging settings and laser power had to be adjusted to account for the increased fluorescence of PAF/GNA resulting in data that appears to show little binding of PAF and GNA. These results provide evidence to suggest that the augmented activity of PAF/GNA is a result of protein anchoring and localisation to the epithelial membrane. Localisation to the stomach region was somewhat surprising as GNA localisation in the *N. lugens* (Brown plant hopper) gut was generally in the intestinal region (posterior midgut) (Powell *et al*., 1998). The observed binding of PAF/GNA was more akin to that observed with mannose-glucose lectin (ConA) was shown to also bind in the stomach region on the *A. pisum* (Sauvion *et al*., 2004). The stomach region of the hemipteran gut is associated with secretion and increased cathepsin activity, and implicated in the digestion of protein, it has also been shown to be slightly more acidic than the remainder of the *A. pisum* gut which may imply an increased proportion of V-ATPase proton pumps (Cristofoletti *et al*., 2003; Ponsen, 1972).

Similar results were demonstrated in Yang *et al*., (2014), the toxicity of amaurobitoxin/GNA was enhanced over its constituent components and both GNA and the fusion protein were shown by fluorescence microscopy to be localised to the *A. pisum* gut and retained after 48 hours of chase feeding with control diet. As Yang *et al*., (2014) performed microscopy on whole insects determination of the region in which PI1a/GNA localised was not possible based on the data presented and the author did not comment further on the topic. Anchoring PAF via the binding of GNA to the gut epithelia may act to facilitate prolonged interactions with V-ATPase pumps, as compared to PAF alone, thus resulting in the observed enhanced toxicity to *A. pisum.*

The LC_{50} values for PAF and PAF/GNA were respectively two-fold and seven-fold higher for *M. persicae* as compared to *A. pisum*. This difference may be attributable to reduced stability of the proteins due to higher proteolytic activity in the gut or salivary secretions of *M. persicae* as compared to *A. pisum*. Previous studies have shown that proteolysis is detrimental to the activity of GNA fusion proteins (Fitches *et al*., 2004; Nakasu *et al*., 2014b). As aphids are known to secrete saliva into the leaves of plants to predigest nutrients the stability of PAF/GNA, PAF and GNA in diet exposed to feeding aphids was assessed. All three proteins appeared similarly stable in both species over 48 hours, with some minor degradation of PAF visible in both species and cleavage of the C-terminal his-tag of PAF/GNA in *M. persicae* treatments. Western analysis of soluble protein extracts prepared from aphids fed on PAF/GNA failed to detect intact PAF/GNA at any time point in either species, suggesting rapid proteolytic cleavage of the fusion protein in both species. The lack of intact PAF/GNA present in the *A. pisum* extracts appears to contradict confocal microscopy results that demonstrate PAF/GNA, and to a lesser extent, GNA and PAF bind to the gut epithelium presumably via interacting either with membrane-bound mannose-glycoproteins or V-type ATPase pumps. However, only the soluble protein fraction was extracted and analysed in the *in vivo* stability assays and it is possible that intact PAF/GNA was present in the insoluble protein fraction that would include membrane bound proteins and thus was not detected.

The intensity of immunoreactive cleavage products in *A. pisum* extracts decreased over time whereas the opposite was observed for *M. persicae*, suggesting that pea aphids consumed less diet over time, and this may be due to antifeedant effects. Further, albeit indirect support for antifeedant effects is suggested by feeding choice experiment results where *A. pisum* showed an increased preference for control over fusion protein diets as compared *M. persicae*.

Unfortunately attempts to evaluate PAF stability *in vivo* were unsuccessful as the histidine tag used for detection of the recombinant protein by western blotting was found to be rapidly cleaved in the gut of both species. *In vivo* experiments showed that GNA was resistant

to proteolysis in both aphid species although susceptibility to cleavage of the C-terminal histidine tag was evidenced by a reduction in mass (by approximately 1 kDa) of the immunoreactive band in blots probed with anti-GNA antibodies and a complete absence of immunoreactivity in blots probed with anti-histidine antibodies, as was seen for PAF. GNA stability and activity in the hemipteran gut has been demonstrated previously (Powell *et al*., 1998). Based on a limited number of previous studies it seems that differences in toxicity of PAF to the different aphid species is unlikely to be due to differences in stability to gut proteolysis. For example Gressent *et al*., (2007), suggested that the degradation of PA1b is unlikely to be the mechanism of resistance to the toxic effects of this peptide, as three extracts from insensitive insects (*Mamestra brassica, Aphis gossypii,* and *Drosophila melanogaster*) and multiple commercial enzymes failed to degrade the protein. Unfortunately, the reason for differential toxicity between species remains unclear as the same study also demonstrated that the epithelial membranes of both sensitive and insensitive species display binding activity with radio labelled PA1b (Gressent *et al*., 2007). The findings suggest the presence of a membrane bound binding site even in insensitive species. The binding site is now known to be the c- subunit of the V_0-ATPase proton pump (Chouabe *et al*., 2011; Muench *et al*., 2014). Further work is required to elucidate the mechanism by which the selectivity of the toxin is conferred.

Previously, Powell *et al*., (1995) demonstrated antifeedant effects of GNA towards the homopteran brown planthopper (*N. lugens*). Initially evidence for suspected antifeedant properties of PAF, GNA and PAF/GNA was sought by weighing feeding chambers to assess the mass of diet consumed but this was unsuccessful due to inconsistent control values. Instead, a free choice assay was performed as an indirect method of assessing antifeedant effects assuming that if aphids preferentially feed on the control diet, it is due to unappealing properties of the test diet. The experimental set up was based on that published by Ghosh *et al*., (2016). When *A. pisum* or *M. persicae* nymphs were provided with a choice of feeding on different diets significantly higher numbers of *A. pisum* were observed on control as compared to PAF, PAF/GNA or GNA containing diets. By contrast, *M. persicae* only exhibited a significant preference for control over PAF/GNA containing diets after 48 hours. Neither species showed a significant preference for diets containing ovalbumin or the control, demonstrating that the effect observed in other conditions is not due to a preference for a low protein diet. The lack of preference in *M. persicae* is possibly due to its generalist feeding strategy; generalists must be more resilient than specialists (such as *A. pisum*) to various plant defence compounds and are more likely to be able to detoxify or sequester deleterious compounds (Ramsey *et al*., 2010).

In the development of any insecticidal compound, it is critical to consider the potential for toxic effects on beneficial insects. Initial chronic feeding assays were performed toward adult honeybees (*Apis melifera*). This preliminary data suggests that PAF/GNA is only mildly toxic to adult honeybees at high doses of PAF/GNA after feeding for 7 days (Data not shown). Previous work with GNA, has demonstrated that the survival of adult honey bees is detrimentally affected by both acute feeding and injection (Nakasu *et al*., 2014). Based on this preliminary data the relatively low level of mortality recorded for honeybees fed on PAF/GNA suggests that at the concentrations required for aphid mortality, there should be little damage to *A. mellifera* populations. The ideal delivery for PAF/GNA to control sap-sucking aphid pests is likely to be through transgenic crops and the concentrations fed to honeybees in this initial study exceed what would feasibly be delivered in the field. More robust and comprehensive testing will be needed to make any claims regarding the effect of PAF/GNA and its components on the survival of adult honeybees. Beyond testing for toxicity, it will be important to examine the effect of PAF/GNA on honeybee memory, as this is key for their survival. Unfortunately, there is no literature reporting the effects of PA1b on honeybees, one statement in a review by Gressent *et al*., (2011) claims *A. mellifera* are insensitive to the insecticidal effects of pea albumin, but no reference is provided or could be found to evidence this claim. Conversely a paper in which pea protein concentrate was used as a protein supplement in *A. mellifera* diet found that it reduced the longevity of bees by 50 % as compared to a pollen-based diet (Nelson, 1976). However, this study was performed nearly ten years before the discovery of pea albumin and the effects of pea protein concentrate on honeybees were not linked to the insecticidal activity of Pa1b (Gatehouse *et al*., 1985; Higgins *et al*., 1986; Nelson, 1976).

Gressent *et al*., (2007) showed substantial deleterious effects on *Harmonia axyridis*, a coleopteran predator of aphids when fed on pea flour. Tritrophic interaction studies will be required to determine if the concentrations of PAF/GNA in aphids feeding on the fusion protein will cause similar effects in *H. axyrids*. However, even with this toxicity, PA1b fusion proteins may have an application in orchard crops and vineyards where *H. axyridis* is considered an invasive pest species (CABI, 2019; Koch, 2003; Pickering *et al*., 2005).

This chapter has provided evidence for the first time that linking recombinant pea albumin to GNA can enhance oral insecticidal activity of PAF, and that this enhancement is facilitated by localisation and retention of the toxin on the apical gut membrane in aphids. Evidence for potential antifeedant properties was also demonstrated. Unfortunately, the mechanism for the differential toxicity between species could not be elucidated. Whilst preliminary data indicates that PAF/GNA shows relatively limited toxicity to honeybees further

studies are required to fully understand the potential impact of the fusion protein upon beneficial insects. Future development of PAF/GNA would require experimentation involving expression in transgenic plants, perhaps under the control of a phloem specific promoter, and this would considerably reduce levels of fusion protein exposure to non-target species. Investigation into whether spraying would be a viable delivery mechanism would also be beneficial as there is some evidence that formulated protein toxins are able to penetrate the insect cuticle (Fanning *et al* 2018). Although spraying would rely on retrograde transport of PAF/GNA from the haemolymph to the gut lumen in order to reach the target site of PA1b. Finally, the reasons underpinning the specificity of action of the pea albumin toxin to different insects is yet to be uncovered and provides an interesting avenue of research.

Chapter 7 |General Discussion

Insect herbivory presents a large threat to food security. Aphids in particular can cause substantial damage to crop yields by transmitting multiple economically important plant viruses and via direct feeding, and could be considered one of the most economically important pests in temperate regions (Blackman and Eastop, 2000; Whitfield *et al*., 2015). The development of insect resistance to chemical controls is a burgeoning issue and there is a pressing need to develop new insecticides which act on novel target sites. Current insecticides tend to effect or inhibit one of five targets in the insect nervous system; namely voltage gated sodium channels, nicotinic acetylcholine receptors, acetylcholinesterase, glutamate receptors and ɣ-aminobutyric (GABA) receptors and chemistries which act on these targets make up three quarters of the worlds pesticide sales (Casida, 2009). The lack of diversity in pesticide target sites of action along with indiscriminate use of insecticides has facilitated the development of resistance towards many families of insecticides (Bass *et al*., 2014; Feyereisen, 1995). Resistance has evolved by multiple mechanisms depending on the target site. Single amino acid mutations in the insecticide binding sites of acetylcholinesterase and the nicotinic acetylcholine receptors have conferred resistance to carbamates and neonicotinoids, respectively (Bass *et al*., 2011; Benting and Nauen, 2004). Mutations away from the insecticide binding site can also cause resistance, for example pyrethroid resistance in *M. persicae* has been associated with mutations in the voltage gated sodium channel as was observed in voltage gated sodium channels in *M. persicae* (Martinez-Torres *et al*., 1999). Further mechanisms include increased expression of detoxifying enzymes such as cytochrome P450 and carboxylesterases through up-regulation and gene duplication leading to resistance against the neonicotinoids and organophosphates, respectively (Devonshire and Moores, 1982; Field *et al*., 1988; Pang *et al*., 2016; Puinean *et al*., 2010). Stacking of these resistance mechanisms can lead to highly resistant strains. For example, highly neonicotinoid resistant *M. persicae* field strains identified in France in 2009 arose as a result of increased expression of cytochrome P450 CYP6CY3 and a mutation in the neonicotinoid binding site on the acetylcholine receptor (Bass *et al*., 2011; Slater *et al*., 2011). Furthermore, thickening of the cuticle can lead to reduced uptake of topically applied compounds (Puinean *et al*., 2010; Strycharz *et al*., 2013). Often resistance to one pesticide can also confer cross resistance, for example the upregulation of carboxylesterase in *M. persicae* in response to exposure to organophosphates has been seen to result in the resistance to both organophosphates and carbamates (Field *et al*., 1988; Liu and Yue, 2000).

The holocyclic life cycle of aphids, which combines parthenogenetic and sexual reproduction, together with short generation times confers these pests with a remarkable

capacity to evolve resistance. *Myzus persicae* has been documented in the Pesticide resistance database, to have evolved field resistance to at least sixteen different active ingredients, which target three different active sites, acetylcholine esterases, voltage gated sodium channels and nicotinic acetylcholine receptors as previously mentioned (Arthropod Pesticide Resistance Database | Michigan State University n.d., Bass *et al.*, 2014).

The successful use of biopesticides, as an alternative to synthetic small molecule pesticides, allows targeting of novel sites in the nervous system such as voltage-gated calcium ion channels or in the insect gut such as V-ATPases, aminopeptidases or cadherin (Bravo *et al*., 2007; Chouabe *et al*., 2011; Tedford *et al*., 2004). The most widely known biopesticides are the Cry toxins from *Bacillus thuringiensis.* The Cry toxins have been applied as organic biopesticides in the form of sprays as early as the 1930s, however their widespread use did not occur until Monsanto commercialised a transgenic line of potato expressing Cry1ac in 1996 (Castagnola and Jurat-Fuentes, 2012). These new insect resistant crops proved to be very effective, providing increased protection against lepidopteran, coleopteran and dipteran pests, greater yields and reduced pesticide usage (Brookes and Barfoot, 2018; Qaim and Zilberman, 2003). As a result the amount of land used to cultivate crops expressing genes from *B. thuringiensis* has seen a 113-fold increase since 1996, with an estimated grown to 101 million hectares globally in 2017 (ISAAA, 2017). However, Bt toxins do not effectively control hemipteran pests, as hemipterans feed on sap there was likely no selective pressure driving *B. thuringiensis* to evolve toxins which act on hemipterans, potentially due to a lack of binding sites and non-optimal conditions for protein activation in the hemipteran gut, however even when processed correctly and bound to the gut low toxicity is observed (Chougule and Bonning, 2012; Lawo *et al*., 2009; Li *et al*., 2011; Porcar *et al*., 2009). The fight against resistance is an ongoing battle from antibiotic resistance of human and livestock pathogens through to the control of insect pests in agriculture. As such new molecules with novel modes of action and ideally acting on novel target sites are desperately needed. This thesis explores the development of *Galanthus nivalis* agglutinin (GNA) fusion protein technology as a potential control strategy for aphid pests, specifically exploiting the small proteinaceous inhibitor of V-ATPases, PA1b.

The presence of an insecticidal compound in pea seeds was first hypothesised by Delobel and Grenier (1993), when they observed toxic effects on weevils of the *Sitophilus* genus fed on split peas. Toxicity was subsequently found to be attributable to the pea seed protein PA1b and its use as an insecticide was patented in France in the 1990s following evidence for toxic effects on *A. pisum* and *E. kuehniella* were demonstrated (Delobel *et al*., 1998). PA1b is a 3.6 kDa, 37 amino acid, *Pisum sativum* seed protein that comprises part of a seed albumin gene

cassette known as PA1 which was initially isolated and described in the 1980s (Gatehouse *et al*., 1985; Higgins *et al*., 1986). The structure of PA1b structure was determined through NMR modelling to be an inhibitor cysteine knot (Jouvensal *et al*., (2003). It has subsequently been shown to reversibly block membrane currents in patch clamping of cultured insect cells (SF9) experiments, in the same manner as bafilomycin (a specific inhibitor of V-ATPases). Inhibition of the V_0 domain of purified V-ATPase proton pumps from *Manduca sexta* has also been demonstrated (Chouabe *et al*., 2011). V-ATPase proton pumps play a key role in the maintenance of gut pH and nutrient transport in insects and inhibition of these pumps explains the entomotoxic properties of PA1b (Wieczorek *et al*., 2009). The V_0 domain of V-ATPase consists of multiple subunits, the specific-subunits which interact with PA1b are *c* and *e* both of which have exposure on the apical side of the gut membrane and are involved in the rotor of V-ATPases (Muench *et al*., 2009, 2014). Recently inhibition of the V-ATPases been shown to cause apoptosis in effected gut cells of the weevil *Sitophilus oryzae* (Eyraud *et al*., 2017). The activity of PA1b has been assessed against species from different orders of insects with variable results, as mentioned previously the literature on activity is limited to two published papers and a patent. Figures in the patent demonstrate toxicity toward *Ephestia kuehniella* and *A. pisum* (Delobel *et al*., 1998). A study by Gressent *et al*., (2007) showed that coleopteran weevils such as *Sitophilus species* and the ladybeetle *Harmonia axyridis* were susceptible to orally delivered PA1b whereas *Tribolium castaneum* was not. Similarly, a significant reduction in the survival of larvae of the mosquito *Culex pipiens* exposed to water containing of 250 ug mL^{-1} of PA1b was recorded after two days. By contrast, PA1b was reported to have no effect on mortality or development when supplied at 1000 µg g^{-1} in artificial diet to lepidopteran *Mamestra brassica*, *Spodotera littoralis* or *Ostrinia nubilalis* larvae (Gressent *et al*., 2007, 2011). Parasitic wasps from the genus *Trichogramma* were found to be unaffected by PA1b in their diet. Gressent *et al*., (2007) also assessed PA1b efficacy in artificial diet assays against three aphid species, finding that *A. pisum* was highly sensitive (500 µg PA1b per mL^{-1}, caused complete mortality after 6 days), whereas greenfly, *Aphis gossypii* was only susceptible at high doses (37 % mortality after 6 days fed on 1000 µg mL^{-1}) and *M. persicae* was insensitive at all doses tested up to 1000 ug mL^{-1}).

This thesis describes work that has significantly enhanced the oral toxicity of pea albumin towards aphids by fusion to snowdrop lectin, GNA. Results reported in Chapter 6 suggest that GNA enhanced the insecticidal activity of pea albumin towards two species of aphid by allowing its retention and localisation to the gut epithelial membrane where V-type ATPase pumps are located. Initial work prior to this project (unpublished) suggested that fusion of PA1b to the N-terminus of GNA resulted in a protein which displayed significantly increased oral

activity toward *A. pisum* as compared to the PA1b activity reported in the literature towards the same pest (Gressent *et al*., 2007). However, these preliminary results were obtained using protein samples of low purity and high (approximately > 95% w/w) carbohydrate contents. In Chapter 4 feeding assays using PA1b/GNA produced previously by Dr Pyati gave similar results to preliminary findings and the correct full-length sequence of recombinant PA1b/GNA was confirmed by LC-MS analysis of digested peptides. However, PA1b/GNA did not agglutinate rabbit erythrocytes, suggesting that the GNA subunit was non-functional (Chapter 4). Increased toxicity in the presence of non-functional GNA would be unexpected as functional GNA has been shown to be necessary for binding to mannose glycans in polypeptides lining the insect gut (Fitches *et al*., 2001). This result led to the hypothesis that perhaps another factor was involved in the extreme toxicity of this fusion protein. Five new expression constructs were created, and proteins were expressed using different fermentation media in an attempt to obtain better quality samples for efficacy tests. In most cases PA1b/GNA fusion proteins were expressed at relatively low levels and histidine affinity chromatography yielded lyophilised samples of poor quality. Nevertheless, a relatively pure sample of PA1b/GNA fusion protein, (designated as Strep/PA1b/GNA) was obtained when purified by strep-tactin affinity chromatography. Purified Strep/PA1b/GNA samples contained only one major contaminating band which was identified as a *P. pastoris* protein with similarities to glucanases. This was determined to be unlikely to affect oral activity as previous work has shown that GNA samples containing this contaminating glucanase show no difference in insecticidal activity as compared to highly purified GNA samples (personal comms – Fitches 2019). The activity of this fusion protein was considerably reduced as compared to PA1b/GNA; feeding of Strep/PA1b/GNA required a final dietary concentration of 0.2 mg mL^{-1} to cause complete mortality in four days as compared to 0.05 mg mL^{-1} required by PA1b/GNA to achieve the same effect. These results further suggested that there may be extraneous factors involved in the initial results obtained for PA1b/GNA. Carry over of imidazole from the purification process would certainly have resulted in increased toxicity as imidazole is known to be toxic to insects (Pence, 1965), but this is unlikely as extensive dialysis is always performed on elution fractions to ensure carry over of this toxic salt does not occur. It is possible that the high carbohydrate content in PA1b/GNA samples may have contributed to the high levels of toxicity with initial samples, for example high sucrose concentration has been shown to severely limit the survival of *A. pisum* on artificial diets (Douglas *et al*., 2006). In addition, previous studies have shown that poor quality protein samples, which contain high levels of carbohydrate and contaminating proteins, are acutely toxic to pea aphids when supplied in diet (unpublished results). This suggests that either the contaminating proteins, high levels of yeast

carbohydrate or a combination of the two are responsible for the acute oral toxicity of PA1b/GNA observed in preliminary experiments.

The expression of PA1b alone in heterologous systems is challenging, a review by Gressent *et al*., (2011) makes reference to personal communication and to a PhD thesis, which was unavailable at the time of writing, indicating that attempts to produce PA1b in both *E. coli* and *P. pastoris* expression hosts were unsuccessful, producing low quantities of non-functional protein. The authors hypothesised that expression of the full PA1 cassette may produce functional protein. Production of functional recombinant leginsulin, a PA1b homologue from soybean, in *E. coli* was achieved using a full length PA1 expression cassette (Yamazaki *et al*., 2003). Transient expression experiments in tobacco have provided further evidence to support this hypothesis as the expression of PA1b alone resulted in between 1.2 -2.4 µg protein per gram fresh leaf as compared to 36.6 or 37.4 µg g^{-1} when the full cassette was expressed (Eyraud *et al*., 2013).

Thus Chapter 5 focused on an alternative production strategy for pea albumin fusion proteins based upon the incorporation of the whole PA1 gene cassette, excluding the signal sequence (PAF). This "full-length" PAF sequence, rather than just the toxin PA1b, was linked to the N-terminus of GNA to create PAF/GNA. A recombinant version of the PA1 cassette alone, designated as PAF was also produced to be used as a control in biological assays, and provides the first evidence of active pea albumin produced in a yeast expression system (Chapter 5 and 6). Expression and purification of these proteins was considerably easier as compared to PA1b/GNA, and samples of greater protein purity (ca. > 90%) and lower levels of carbohydrate were produced using a range of native purification strategies. Various optimal conditions were observed, PAF showed better expression and recovery from fermentations maintained at pH 5, as compared to pH 4.5. By contrast, highest expression levels for PAF/GNA were recorded for fermentations conducted at pH 4.5 but the recovery of highly pure protein samples from nickel matrices was poor. Based on findings reported in Chapter 3, where increased binding of GNA to nickel columns was demonstrated in supernatants containing glucose as compared to glycerol, PAF/GNA was also expressed using glucose as an alternative carbon source to glycerol. Whilst using glucose appeared to reduce PAF/GNA expression levels (as compared to fermentation with glycerol), there was (as observed for GNA in Chapter 3) an increase in the amount of relatively pure protein recovered. Little literature exists to explain the difference in binding of proteins to nickel matrices in the presence of glycerol or glucose but it was concluded that glycerol may interfere with complex formation (Ellis *et al*., 1989).

Biochemical characterisation of PAF and PAF/GNA was performed with LC-MS analysis of digested samples confirming the presence of intact and correct amino acid sequences for both recombinant proteins. Glycosylation of recombinant PAF and PAF/GNA when expressed in *P. pastoris* was analysed using Schiff-periodic acid and no substantial differences were observed in the degree of glycosylation of proteins produced at different fermentation pHs or using different carbon sources. Unlike initial PA1b/GNA samples, the ability of PAF and PAF/GNA to cross link rabbit erythrocytes was confirmed through hemagglutination assays.

The insecticidal activity of recombinant PAF and PAF/GNA was evaluated in Chapter 6 with results confirming that linkage of PAF to GNA resulted in significantly enhanced activity of PAF towards *A. pisum* and *M. persicae*. Experiments were then performed in an attempt to understand the mechanisms underpinning the enhanced toxicity of the fusion protein as compared to PAF alone. On a molar basis recombinant PAF showed somewhat comparable toxicity towards *A. pisum* to that of PA1b purified from pea seeds as described previously by Gressent *et al*., (2007). The four-fold increased LC_{50} for PAF (112 nmol mL^{-1}) against *M. persicae* as compared to *A. pisum* could reflect the insensitivity *M. persicae* to purified PA1b reported in the literature. Fusion of PAF to the N-terminus of GNA resulted in significantly increased toxicity towards both *A. pisum* and *M. persicae* with PAF/GNA LC_{50}s respectively, ca. 12-fold and 4-fold lower as compared to PAF alone. *Sitobion avenae* strangely showed no significant difference in PAF and PAF/GNA LC_{50} values. Linking protein-based insecticidal neurotoxins to GNA has previously been shown to enhance oral activity against a range of insect pests (Fitches *et al*., 2002, 2004, 2012; Nakasu *et al*., 2014a; Powell *et al*., 2019; Trung *et al*., 2006; Yang, 2015; Yang *et al*., 2014). All toxins previously fused to GNA have target sites in the insect nervous system or haemolymph and, as such, GNA has been utilised as a carrier to transport attached toxins across the gut epithelium, into the haemolymph and potentially localise the toxin in the nervous system (Fitches *et al*., 2001, 2012). However, pea albumin – GNA fusion proteins are unique in that PAF targets, V-ATPases present on the apical membrane of the alimentary canal and this thesis provides the first evidence that GNA can enhance the oral activity of a toxin with a target site within the gut lumen. Retention in the gut was thought to be key to the increased efficacy of PAF/GNA and this was demonstrated using confocal microscopy of aphid guts extracted from aphids fed on fluorescently labelled proteins PAF/GNA or PAF or GNA. When aphids were chase-fed control diets, PAF/GNA was retained in greater quantities than either PAF or GNA alone in the gut of *A. pisum.* In agreement with aphid bioassay results, linking PAF to GNA appeared to have synergistic rather than purely additive effects, as the fluorescence for PAF/GNA was greater than the sum of PAF and GNA. That GNA and GNA based fusion proteins bind to the

hemipteran gut has been previously demonstrated (Powell *et al*., 1998; Yang *et al*., 2014). In fact, GNA has been shown to bind to *A. pisum* membrane-bound aminopeptidase, which makes up 15.6 % of *A. pisum* gut protein (Cristofoletti *et al*., 2006). Previous GNA based fusion proteins have shown localisation to the aphid gut and PA1b has displayed binding affinity for multiple insect gut membrane extracts (Gressent *et al*., 2007; Yang *et al*., 2014).

Susceptibility of different aphid species to the effects of insecticidal proteins is known to be variable, with specialist species (such as *A. pisum*) generally exhibiting greater susceptibility as compared to generalists (such as *M. persicae*). GNA is no exception to this with *M. persicae* requiring a 10-fold higher concentration of GNA than *A. pisum* to cause 50 % reductions in survival (Rahbé *et al*., 1995b; Sadeghi *et al*., 2009; Sauvion *et al*., 1996; Zapata *et al*., 2016). Differences in susceptibility to insecticidal proteins have been generally attributed to differences in proteolytic activity within the insect gut; *A. pisum* has relatively low levels of proteolytic activity and detoxification enzymes as compared to *M. persicae.* The increased proteolytic activity in *M. persicae* is likely to be an adaptive response to feeding on high protein diet, rapid transcriptional changes, including upregulation of proteases, have been documented when subjecting *M. persicae* to different diets in series (Mathers *et al*., 2017; Ramsey *et al*., 2010). PAF, GNA and PAF/GNA were all observed to be stable in the presence of salivary secretions which are known to contain proteases from both species (Furch *et al*., 2015). It was therefore a surprise that no intact fusion protein was detected *in vivo* in total protein extracts from pea aphids fed on diets containing PAF/GNA. The small quantities of diet, and therefore protein consumed by aphids in combination with antifeedant effects makes the detection of protein challenging as aphids must be left to feed for relatively long periods of time. Cleavage of the fusion protein in *M. persicae* was expected as this has been documented previously for Hv1a/GNA (Nakasu *et al*., 2014b). The lack of intact PAF/GNA in *A. pisum* extracts was surprising as they are reported to have low proteolytic activity in the gut and previous work with Hv1a/GNA and *S. avenae* another specialist species showed intact fusion protein after 24 hours of feeding (Nakasu *et al*., 2014b; Rahbé *et al*., 1995b). Ambiguity in the methods of Nakasu *et al*., (2014) make direct comparison difficult as *S. avenae* may have been fed at the same dose or at a concentration approximately double to PAF/GNA in this study. It was also observed during the feeding stage of the stability assays that several the *A. pisum* were not feeding, and this may well have influenced the result. Furthermore, as microscopy experiments suggested that PAF/GNA binds to the gut epithelial membrane it may be that intact PAF/GNA was not present in the soluble fraction, extracted in PBS, used in the *in vivo* analysis.

Consequently, attempts were made to assess potential antifeedant effects of PAF/GNA and its components in both *A. pisum* and *M. persicae,* by using a free choice assay, where aphids were given two diet options one containing the test protein one without. Nymphs of *A. pisum* displayed significant preference for diets without the toxins present, whereas *M. persicae* showed no preference, except for PAF/GNA after 48 hours feeding. The inclusion of ovalbumin as a control, suggests that the aphids are averse to PAF, PAF/GNA and GNA rather than purely adverse to protein rich diets. Antifeedant effects of GNA based fusion proteins towards hemipteran pests have not been directly demonstrated previously although antifeedant effects of GNA alone on brown plant hopper (*Nilaparvata lugens*) have previously been documented (Powell *et al*., 1995). Whether antifeedant effects are advantageous or deleterious for an insecticidal compound is dependent on the delivery mechanism, if PAF/GNA was expressed in a transgenic crop, antifeedant effects may serve to protect crops from damage by driving the pests away to refuges field borders or trap crops.

The assessment of any novel insecticide against non-target beneficial species is essential. Of note is that the outdoor use of three neonicotinoid pesticides has recently been banned in the EU due to their adverse effects on pollinators (European Commission, 2018a, 2018b, 2018c; Woodcock *et al*., 2017). Consequently, further work will be required to assess the effect of PAF/GNA was assessed against the European honeybee (*Apis mellifera*). Initial assays suggest that PAF/GNA was only mildly toxic when supplied to adult *A. mellifera* in a chronic feeding assay. GNA and Hv1a/GNA fusion proteins have previously been shown to be non-toxic to *A. mellifera* adults in feeding and injection studies (Nakasu *et al*., 2014a; Powell *et al*., 2019). However, further studies are required to understand if PAF or PAF/GNA are indeed toxic to honeybees. PA1b has been demonstrated to have detrimental effects on beneficial insect *H. axyridis*, and tri-trophic interaction studies would need to be performed to determine if this effect would be detrimental in a field scenario (Gressent *et al*. 2007).

The potential for commercial application of PAF/GNA in the field is limited by the target specificity of PA1b, more specifically its lack of toxicity towards lepidopteran pests (Gressent *et al*., 2007, 2011). PAF/GNA may find a place in mosquito control, as PA1b shows activity against *C. pipens* larvae, the vector of Japanese encephalitis and the West Nile virus. PA1b has also been demonstrated cause significant detrimental effects on the larval development and survival of *H. axyridis* which may limit the application of PAF/GNA, although further testing is required. *Harmonia axyridis* is generally considered beneficial due to its predation of aphids (Gressent *et al*., 2007). However it is considered a pest of orchard fruit in the USA or in vineyards where it is considered undesirable due its adverse effects on flavour of the resulting wine presenting a

potential niche for PAF/GNA to fill (CABI, 2019; Koch, 2003; Pickering *et al*., 2005). One further potential application of this fusion protein is in the protection of dry goods such as flour, multiple weevils of the *Sitophilus* genus are sensitive to PA1b and PAF/GNA should be tested to determine if it as with both *A. pisum* and *M. persicae* has enhance insecticidal activity (Gressent *et al*., 2003). Application as a spray is unlikely due the toxicity of PA1b relying on oral rather than topical delivery. However, recent evidence suggests that with correct formulation topically applied neurotoxic proteins can penetrate across the insect cuticle to the haemolymph (Fanning *et al*., 2018), and it is known that GNA, when injected into lepidopteran larvae traverses the gut epithelia from the basal to the apical membrane (unpublished results). Thus, it is possible that topical application of PAF/GNA may facilitate delivery of PAF to the aphid gut epithelia although this requires further investigation and would rely on the retrograde transport of the toxin from the haemolymph to the gut. The most feasible delivery route for PAF/GNA in the field is via phloem expression *in planta*, where it has potential use as a strategy to control hemipteran pests, potentially filling the gap left by the relative infectiveness Bt Cry toxins towards these pests (Chougule and Bonning, 2012; Frankenhuyzen, 2009). Expression in the plant phloem would also reduce the exposure of non-target species such as *Apis mellifera*. Gene pyramiding with a Bt expressing crop could provide protection against a larger number of pests that PAF/GNA could achieve on its own. Hv1a/GNA when expressed in transgenic Arabidopsis has been shown to confer a degree of resistance to *M. persicae* with two different lines of plants causing 19.6 % and 37 % mortality relative to the negative control (Nakasu *et al*., 2014b). Comparison of the activity of Hv1a/GNA in artificial diet feeding assays from the same study to PAF/GNA at an equivalent molar dose suggests that these proteins have similar efficacy towards *M. persicae* with Hv1a/GNA causing 90% mortality after 8 days, as compared to 100 % when fed PAF/GNA. PAF/GNA may be particularly effective in transgenic legumes as PAF/GNA has at least a 2 fold greater effect on *A. pisum* than Pro-Hv1a/GNA although this would need to be tested *in planta* (Yang, 2015). Even if expressed *in planta* the issue of resistance remains present, this could perhaps be avoided through gene stacking with another fusion protein or aphidicidal protein which acts on a separate target site such as Hv1a/GNA or PI1a/GNA (Nakasu *et al*., 2014b; Yang *et al*., 2014). Further work is required to establish if PAF/GNA can be expressed *in planta* as an intact functional fusion protein without detrimental effects upon plant development, as reported previously for of the PA1b soybean homologue, leginuslin, which was shown to stimulate cell proliferation in cultured carrot cells (Yamazaki *et al*., 2003). Recently it has also come to light that perhaps there are other Ab1 peptides from legumes which may have similar or greater insecticidal effects such as a homologue to PA1b from *Medicago truncatula*

(Royer *et al*., 2018). These homologues will be worth investigating in the context of fusion proteins as they may provide a more suitable compound with greater entomotoxic activity or other benefits which may make them more applicable to field application.

Additionally, this thesis documents the first evidence that fusion of GNA with insecticidal proteins with a target site in the insect gut can enhance the oral activity through allowing localisation to the gut epithelial wall. This opens up possibilities for the development of other novel fusion proteins utilising insecticidal compounds other than the previously explored neuropeptides. Perhaps fusion to a plant protease inhibitor, or one of the recently discovered Ab1 homologues mentioned above, which have been demonstrated to have some insecticidal activity, may result in enhanced aphidical activity as aphid digestive enzymes are often localized to the epithelial membrane (Cristofoletti *et al*., 2003; Rahbé *et al*., 2003).

The data present in this thesis also serves to further highlight the potential of targeting V-ATPase pumps as a novel target for both pesticides and biopesticides. These pumps serve a key role in insect gut homeostasis and nutrient uptake (O'Donnell, 2017). Disruption of these functions could represent an effective novel and specific target for the control of many insect pests. Unfortunately the known chemical inhibitors of V-ATPase are limited in their application due to their toxicity profile (Li *et al*., 2017). Perhaps the utilisation of PAF/GNA in conjunction with RNAi designed to knock down the expression of V-ATPase may result in synergistic effects and should be tested in the future.

Overall, this study has shown that PAF (PA1b+PA1a) to GNA has significantly enhanced its activity against two aphid species. Further, evidence is provided to suggest that this enhancement is due to GNA allowing the localisation of the PA1b toxin to the gut epithelium. This has provided the foundations for many lines of future research as outlined above. This is important because novel, safe and effective pesticides will be required in order to feed the ever-increasing global population.

References:

Abdourahime, H., Anastassiadou, M., Arena, M., Auteri, D., Barmaz, S., Brancato, A., Brocca, D., Bura, L., Cabrera, L.C., Chiusolo, A., *et al*. (2019). Peer review of the pesticide risk assessment of the active substance thiacloprid. EFSA J. *17*, e05595.

Ahmad, M., Hirz, M., Pichler, H., and Schwab, H. (2014). Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. Appl. Microbiol. Biotechnol. *98*, 5301–5317.

Aktar, W., Sengupta, D., and Chowdhury, A. (2009). Impact of pesticides use in agriculture: their benefits and hazards. Interdiscip. Toxicol. *2*, 1–12.

Almeida, M.S., Cabral, K.S., Neves de Medeiros, L., Valente, A.P., Almeida, F.C.L., and Kurtenbach, E. (2001). cDNA Cloning and Heterologous Expression of Functional Cysteine-Rich Antifungal Protein Psd1 in the Yeast Pichia pastoris. Arch. Biochem. Biophys. *395*, 199–207.

Bai, Y., and Shen, W.-C. (2006). Improving the Oral Efficacy of Recombinant Granulocyte Colony-Stimulating Factor and Transferrin Fusion Protein by Spacer Optimization. Pharm. Res. *23*, 2116–2121.

Barre, A., Bourne, Y., Van Damme, E.J.M., Peumans, W.J., and Rougé, P. (2001). Mannose-binding plant lectins: Different structural scaffolds for a common sugar-recognition process. Biochimie *83*, 645–651.

Basnet, S., and Kamble, S.T. (2018). RNAi-Mediated Knockdown of vATPase Subunits Affects Survival and Reproduction of Bed Bugs (Hemiptera: Cimicidae). J. Med. Entomol. *55*, 540–546.

Bass, C., Puinean, A.M., Andrews, M., Cutler, P., Daniels, M., Elias, J., Paul, V.L., Crossthwaite, A.J., Denholm, I., Field, L.M., *et al*. (2011). Mutation of a nicotinic acetylcholine receptor β subunit is associated with resistance to neonicotinoid insecticides in the aphid Myzus persicae. BMC Neurosci. *12*.

Bass, C., Puinean, A.M., Zimmer, C.T., Denholm, I., Field, L.M., Foster, S.P., Gutbrod, O., Nauen, R., Slater, R., and Williamson, M.S. (2014). The evolution of insecticide resistance in the peach potato aphid, Myzus persicae. Insect Biochem. Mol. Biol. *51*, 41–51.

Baum, J.A., Bogaert, T., Clinton, W., Heck, G.R., Feldmann, P., Ilagan, O., Johnson, S., Plaetinck, G., Munyikwa, T., Pleau, M., *et al*. (2007). Control of coleopteran insect pests through RNA interference. Nat. Biotechnol. *25*, 1322–1326.

Baumgartner, P., Harper, K., Raemaekers, R.J.M., Durieux, A., Gatehouse, A.M.R., Davies, H.V., and Taylor, M.A. (2003). Large-scale production and purification of recombinant *Galanthus nivalis* agglutinin (GNA) expressed in the methylotrophic yeast *Pichia pastoris.* Biotechnol. Lett. *25*, 1281–1285.

Beard, J. (2006). DDT and human health. Sci. Total Environ. *355*, 78–89.

van Bel, A.J.E., and Will, T. (2016). Functional Evaluation of Proteins in Watery and Gel Saliva of Aphids. Front. Plant Sci. *07*.

Benting, J., and Nauen, R. (2004). Biochemical evidence that an S431F mutation in acetylcholinesterase-1 ofAphis gossypii mediates resistance to pirimicarb and omethoate. Pest Manag. Sci. *60*, 1051–1055.

Beyenbach, K.W., and Wieczorek, H. (2006). The V-type H+ ATPase: molecular structure and function, physiological roles and regulation. J. Exp. Biol. *209*, 577–589.

Blackman, R., and Eastop, V. (2000). Aphids on the Worlds Crops: An Identification and Information Guide John Wiley and Sons. N. Y.

Blackman, R.L., and Eastop, V.F. (2007). Taxonomic issues. Aphids Crop Pests 1–29.

Bonning, B.C., and Chougule, N.P. (2014). Delivery of intrahemocoelic peptides for insect pest management. Trends Biotechnol. *32*, 91–98.

Bornhorst, J.A., and Falke, J.J. (2000). Purification of proteins using polyhistidine affinity tags. Methods Enzymol. *326*, 245–254.

Bowman, E.J., Siebers, A., and Altendorf, K. (1988). Bafilomycins: a class of inhibitors of membrane ATPases from microorganisms, animal cells, and plant cells. Proc. Natl. Acad. Sci. *85*, 7972–7976.

Brake, A.J., Merryweather, J.P., Coit, D.G., Heberlein, U.A., Masiarz, F.R., Mullenbach, G.T., Urdea, M.S., Valenzuela, P., and Barr, P.J. (1984). Alpha-factor-directed synthesis and secretion of mature foreign proteins in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *81*, 4642–4646.

Bravo, A., Gill, S.S., and Soberon, M. (2007). Mode of action of Bacillus thuringiensis Cry and Cyt toxins and their potential for insect control. Toxicon *49*, 423–435.

Bravo, A., Likitvivatanavong, S., Gill, S.S., and Soberón, M. (2011). Bacillus thuringiensis: A story of a successful bioinsecticide. Insect Biochem. Mol. Biol. *41*, 423–431.

Brookes, G., and Barfoot, P. (2018). Environmental impacts of genetically modified (GM) crop use 1996-2016: Impacts on pesticide use and carbon emissions. GM Crops Food *9*, 109–139.

CABI (2018a). Myzus persicae (green peach aphid).

CABI (2018b). Acyrthosiphon pisum (pea aphid).

CABI (2018c). Sitobion avenae (wheat aphid).

CABI (2019). *Harmonia axyridis.* InL Invasive Species Compendium, Wallingford, UK: CAB International.

Çalık, P., Bayraktar, E., İnankur, B., Soyaslan, E.Ş., Şahin, M., Taşpınar, H., Açık, E., Yılmaz, R., and Özdamar, T.H. (2010). Influence of pH on recombinant human growth hormone production by Pichia pastoris. J. Chem. Technol. Biotechnol. *85*, 1628–1635.

Casida, J.E. (2009). Pest Toxicology: The Primary Mechanisms of Pesticide Action. Chem. Res. Toxicol. *22*, 609–619.

Castagnola, A.S., and Jurat-Fuentes, J.L. (2012). Bt Crops: Past and Future. In Bacillus Thuringiensis Biotechnology, E. Sansinenea, ed. (Dordrecht: Springer Netherlands), pp. 283–304.

Chang, S.-W., Shieh, C.-J., Lee, G.-C., Akoh, C.C., and Shaw, J.-F. (2006). Optimized Growth Kinetics of *Pichia pastoris* and Recombinant *Candida rugosa* LIP1 Production by RSM. J. Mol. Microbiol. Biotechnol. *11*, 28–40.

Chouabe, C., Eyraud, V., Da Silva, P., Rahioui, I., Royer, C., Soulage, C., Bonvallet, R., Huss, M., and Gressent, F. (2011). New Mode of Action for a Knottin Protein Bioinsecticide: Pea Albumin 1 Subunit b (PA1b) is the first peptidic inhibitor of V-ATPase. J. Biol. Chem. *286*, 36291–36296.

Chougule, N.P., and Bonning, B.C. (2012). Toxins for Transgenic Resistance to Hemipteran Pests. Toxins *4*, 405–429.

Chougule, N.P., Li, H., Liu, S., Linz, L.B., Narva, K.E., Meade, T., and Bonning, B.C. (2013). Retargeting of the Bacillus thuringiensis toxin Cyt2Aa against hemipteran insect pests. Proc. Natl. Acad. Sci. *110*, 8465–8470.

Cranston, P.S., and Gullan, P.J. (2009). Chapter 199 - Phylogeny of Insects. In Encyclopedia of Insects (Second Edition), V.H. Resh, and R.T. Cardé, eds. (San Diego: Academic Press), pp. 780–793.

Cregg, J.M., Cereghino, J.L., Shi, J., and Higgins, D.R. (2000). Recombinant Protein Expression in Pichia pastoris. Mol. Biotechnol. *16*, 23–52.

Crickmore, N., Zeigler, D.R., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J., Bravo, A., and Dean, D.H. (2014). Bacillus thuringiensis toxin nomenclature. Bacillus Thuringiensis.

Cristofoletti, P.T., Ribeiro, A.F., Deraison, C., Rahbé, Y., and Terra, W.R. (2003). Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid Acyrthosiphon pisum. J. Insect Physiol. *49*, 11–24.

Cristofoletti, P.T., Mendonça de Sousa, F.A., Rahbé, Y., and Terra, W.R. (2006). Characterization of a membrane-bound aminopeptidase purified from Acyrthosiphon pisum midgut cells: A major binding site for toxic mannose lectins. FEBS J. *273*, 5574–5588.

Da Silva, P., Rahioui, I., Laugier, C., Jouvensal, L., Meudal, H., Chouabe, C., Delmas, A.F., and Gressent, F. (2010). Molecular Requirements for the Insecticidal Activity of the Plant Peptide Pea Albumin 1 Subunit b (PA1b). J. Biol. Chem. *285*, 32689–32694.

Daly, N.L., and Craik, D.J. (2011). Bioactive cystine knot proteins. Curr. Opin. Chem. Biol. *15*, 362–368.

D'arcy, C.J. (1995). Symptomatology and host range of barley yellow dwarf. Barley Yellow Dwarf *40*, 9–28.

DeBach, P., and Rosen, D. (1991). Biological control by natural enemies (Cambridge, [England] ; New York: Cambridge University Press).

Dedryver, C.-A., Le Ralec, A., and Fabre, F. (2010). The conflicting relationships between aphids and men: A review of aphid damage and control strategies. Pucerons Modèles Biol. Ravag. Cult. *333*, 539–553.

Delobel, B., and Grenier, A.M. (1993). Effect of non-cereal food on cereal weevils and tamarind pod weevil (Coleoptera: Curculionidae). J. Stored Prod. Res. *29*, 7–14.

Delobel, B., Grenier, A., Gueguen, J., Ferrasson, E., and Mbaiguinam, M. (1998). Utilisation d'un polypeptide dérivé d'une albumine PA1b de légumineuse comme insecticide. Fr. Pat. *98*, 05877.

Devonshire, A.L., and Moores, G.D. (1982). A carboxylesterase with broad substrate specificity causes organophosphorus, carbamate and pyrethroid resistance in peach-potato aphids (Myzus persicae). Pestic. Biochem. Physiol. *18*, 235–246.

Douglas, A.E. (2006). Phloem-sap feeding by animals: problems and solutions. J. Exp. Bot. *57*, 747–754.

Douglas, A.E., Price, D.R., Minto, L.B., Jones, E., Pescod, K.V., Francois, C.L., and Boonham, N. (2006). Sweet problems: insect traits defining the limits to dietary sugar utilisation by the pea aphid, Acyrthosiphon pisum. J. Exp. Biol. *209*, 1395–1403.

Dow (1992). pH Gradients in the Lepidopteran Midgut. J. Exp. Biol. *172*, 355–375.

Down, R.E., Fitches, E.C., Wiles, D.P., Corti, P., Bell, H.A., Gatehouse, J.A., and Edwards, J.P. (2006). Insecticidal spider venom toxin fused to snowdrop lectin is toxic to the peach-potato aphid,Myzus persicae (Hemiptera: Aphididae) and the rice brown planthopper,Nilaparvata lugens (Hemiptera: Delphacidae). Pest Manag. Sci. *62*, 77–85.

Duan, J.J., Marvier, M., Huesing, J., Dively, G., and Huang, Z.Y. (2008). A Meta-Analysis of Effects of Bt Crops on Honey Bees (Hymenoptera: Apidae). PLoS ONE *3*, e1415.

Ealing, P.M., Hancock, K.R., and White, D.W.R. (1994). Expression of the pea albumin 1 gene in transgenic white clover and tobacco. Transgenic Res. *3*, 344–354.

Eissazadeh, S., Moeini, H., Dezfouli, M.G., Heidary, S., Nelofer, R., and Abdullah, M.P. (2017). Production of recombinant human epidermal growth factor in Pichia pastoris. Braz. J. Microbiol. *48*, 286–293.

Elbetieha, A., Da'as, S.I., Khamas, W., and Darmani, H. (2001). Evaluation of the Toxic Potentials of Cypermethrin Pesticide on Some Reproductive and Fertility Parameters in the Male Rats. Arch. Environ. Contam. Toxicol. *41*, 522–528.

Ellis, N.A., Hubbard, C.D., Shirodkar, S.M., Guardado, P., Moya, M.-L., Sanchez-Burgos, F., and Burgess, J. (1989). Medium effects upon the kinetics of formation of nickel(II) and cobalt(II) pyridine 2-azo-p-dimethylaniline. Transit. Met. Chem. *14*, 466–470.

Elzinga, D.A., De Vos, M., and Jander, G. (2014). Suppression of Plant Defenses by a *Myzus persicae* (Green Peach Aphid) Salivary Effector Protein. Mol. Plant. Microbe Interact. *27*, 747–756.

van Emden, H.F., and Harrington, R. (2007). Aphids as Crop Pests (CABI).

van Emden, H.F., Eastop, V.F., Hughes, R.D., and Way, M.J. (1969). The Ecology of *Myzus persicae*. Annu. Rev. Entomol. *14*, 197–270.

European Commission (2017). Integrated Pest Management (IPM).

European Commission (2018a). Commission Implementing Regulation (EU) 2018/783 of 29 May 2018 amending Implementing Regulation (EU) No 540/2011 as regards the conditions of approval of the active substance imidacloprid (Text with EEA relevance.).

European Commission (2018b). Commission Implementing Regulation (EU) 2018/784 of 29 May 2018 amending Implementing Regulation (EU) No 540/2011 as regards the conditions of approval of the active substance clothianidin (Text with EEA relevance.).

European Commission (2018c). Commission Implementing Regulation (EU) 2018/785 of 29 May 2018 amending Implementing Regulation (EU) No 540/2011 as regards the conditions of approval of the active substance thiamethoxam (Text with EEA relevance.).

Eyraud, V., Karaki, L., Rahioui, I., Sivignon, C., Da Silva, P., Rahbé, Y., Royer, C., and Gressent, F. (2013). Expression and Biological Activity of the Cystine Knot Bioinsecticide PA1b (Pea Albumin 1 Subunit b). PLoS ONE *8*, e81619.

Eyraud, V., Balmand, S., Karaki, L., Rahioui, I., Sivignon, C., Delmas, A.F., Royer, C., Rahbé, Y., Da Silva, P., and Gressent, F. (2017). The interaction of the bioinsecticide PA1b (Pea Albumin 1 subunit b) with the insect V-ATPase triggers apoptosis. Sci. Rep. *7*.

Fang, J., and Ewald, D. (2004). Expression cloned cDNA for 10-deacetylbaccatin III-10-O-acetyltransferase in Escherichia coli: a comparative study of three fusion systems. Protein Expr. Purif. *35*, 17–24.

Fanning, P.D., VanWoerkom, A., Wise, J.C., and Isaacs, R. (2018). Assessment of a commercial spider venom peptide against spotted-wing Drosophila and interaction with adjuvants. J. Pest Sci. *91*, 1279–1290.

FAO (2011). The State of the World's Land and Water Resources for Food and Agriculture 2011, Managing systems at risk (Food and Agriculture Organisation of the United Nations).

FAO (2013). Official food and agricultural organisation of the united nations figures.

FAO (2016). The State of Food and Agriculture 2016, Climage Change, Agriculture and Food Security (Food and Agriculture Organisation of the United Nations).

FAO (2019). Plant Production and Protection Division: Integrated Pest Management.

FAO, IFAD, UNICEF, WFP, and WHO (2019). The State of Food Security and Nutrition in the World 2019. Safeguardig against economic slowdows and downturns. (Rome).

Ferry, N., and Gatehouse, A.M.R. (2010). Transgenic crop plants for resistance to biotic stress. In Biotech Crops, Concepts, Construction, Contribution and Concerns, C. Kole, C. Michler, A. Abbott, and T. Hall, eds. (Dodrecht: Spriner), pp. 1–65.

Feyereisen, R. (1995). Molecular biology of insecticide resistance. Toxicol. Lett. *82–83*, 83–90.

Field, L.M., Devonshire, A.L., and Forde, B.G. (1988). Molecular evidence that insecticide resistance in peach-potato aphids (*Myzus persicae* Sulz.) results from amplification of an esterase gene. Biochem. J. *251*, 309–312.

Fitches, E.C., and Gatehouse, J.A. (1998). A comparison of the short and long term effects of insecticidal lectins on the activities of soluble and brush border enzymes of tomato moth larvae (Lacanobia oleracea). J. Insect Physiol. *44*, 1213–1224.

Fitches, E.C., Gatehouse, A.M.R., and Gatehouse, J.A. (1997). Effects of snowdrop lectin (GNA) delivered via artificial diet and transgenic plants on the development of tomato moth (Lacanobia oleracea) larvae in laboratory and glasshouse trials. J. Insect Physiol. *43*, 727–739.

Fitches, E.C., Woodhouse, S.D., Edwards, J.P., and Gatehouse, J.A. (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.C., Audsley, N., Gatehouse, J.A., and Edwards, J.P. (2002). Fusion proteins containing neuropeptides as novel insect contol agents: snowdrop lectin delivers fused allatostatin to insect haemolymph following oral ingestion. Insect Biochem. Mol. Biol. *32*, 1653–1661.

Fitches, E.C., Edwards, M.G., Mee, C., Grishin, E., Gatehouse, A.M.R., Edwards, J.P., and Gatehouse, J.A. (2004). 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.C., Pyati, P., King, G.F., and Gatehouse, J.A. (2012). Fusion to snowdrop lectin magnifies the oral activity of insecticidal ω-Hexatoxin-Hv1a peptide by enabling its delivery to the central nervous system. PloS One *7*, e39389.

Fletcher, J.I., Smith, R., O'Donoghue, S.I., Nilges, M., Connor, M., Howden, M.E., Christie, M.J., and King, G.F. (1997). The structure of a novel insecticidal neurotoxin, omega-atracotoxin-HV1, from the venom of an Australian funnel web spider. Nat. Struct. Biol. *4*, 559–566.

Forgac, M. (2007). Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology. Nat. Rev. Mol. Cell Biol. *8*, 917–929.

Frankenhuyzen, K. van (2009). Insecticidal activity of Bacillus thuringiensis crystal proteins. J. Invertebr. Pathol. *101*, 1–16.

Freire, C., and Koifman, S. (2012). Pesticide exposure and Parkinson's disease: Epidemiological evidence of association. NeuroToxicology *33*, 947–971.

Fu, K.-Y., Guo, W.-C., Lü, F., Liu, X., and Li, G.-Q. (2014). Response of the vacuolar ATPase subunit E to RNA interference and four chemical pesticides in Leptinotarsa decemlineata (Say). Pestic. Biochem. Physiol. *114*, 16–23.

Furch, A.C.U., van Bel, A.J.E., and Will, T. (2015). Aphid salivary proteases are capable of degrading sieve-tube proteins. J. Exp. Bot. *66*, 533–539.

Gatehouse, A.M.R., Powell, K.S., Peumans, W.J., Van Damme, E.J., and Gatehouse, J.A. (1995). Insecticidal properties of plant lectins: their potential in plant protection. Lectins Biomed. Perspect. 35–57.

Gatehouse, A.M.R., Down, R.E., Powell, K.S., Sauvion, N., Rahbé, Y., Newell, C.A., Merryweather, A., Hamilton, W.D., and Gatehouse, J.A. (1996). Transgenic potato plants with enhanced resistance to the peach-potato aphid Myzus persicae. Entomol. Exp. Appl. *79*, 295–307.

Gatehouse, A.M.R., Davison, G.M., Newell, C.A., Merryweather, A., Hamilton, W.D., Burgess, E.P., Gilbert, R.J., and Gatehouse, J.A. (1997). Transgenic potato plants with enhanced resistance to the tomato moth, Lacanobia oleracea: growth room trials. Mol. Breed. *3*, 49–63.

Gatehouse, J.A., Gilroy, J., Hoque, M.S., and Croy, R.R. (1985). Purification, properties and amino acid sequence of a low-Mr abundant seed protein from pea (Pisum sativum L.). Biochem. J. *225*, 239–247.

Geiger, F., Bengtsson, J., Berendse, F., Weisser, W.W., Emmerson, M., Morales, M.B., Ceryngier, P., Liira, J., Tscharntke, T., Winqvist, C., *et al*. (2010). Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland. Basic Appl. Ecol. *11*, 97–105.

Ghosh, A., Das, A., Vijayanandraj, S., and Mandal, B. (2016). Cardamom Bushy Dwarf Virus Infection in Large Cardamom Alters Plant Selection Preference, Life Stages, and Fecundity of Aphid Vector, *Micromyzus kalimpongensis* (Hemiptera: Aphididae). Environ. Entomol. *45*, 178–184.

Giri, S., Giri, A., Sharma, G.D., and Prasad, S.B. (2002). Mutagenic effects of carbosulfan, a carbamate pesticide. Mutat. Res. Toxicol. Environ. Mutagen. *519*, 75–82.

Gómez, I., Sánchez, J., Miranda, R., Bravo, A., and Soberón, M. (2002). Cadherin-like receptor binding facilitates proteolytic cleavage of helix alpha-1 in domain I and oligomer pre-pore formation of Bacillus thuringiensis Cry1Ab toxin. FEBS Lett. *513*, 242–246.

Goulson, D. (2014). Pesticides linked to bird declines. Nature *511*, 295–296.

Grabe, M., Wang, H., and Oster, G. (2000). The Mechanochemistry of V-ATPase Proton Pumps. Biophys. J. *78*, 2798–2813.

Gray, S.M. (2008). Aphid Transmission of Plant Viruses. Curr. Protoc. Microbiol. *10*, 16B.1.1-16B.1.10.

Grenier, A.-M., Mbaiguinam, M., and Delobel, B. (1997). Genetical analysis of the ability of the rice weevil Sitophilus oryzae (Coleoptera, Curculionidae) to breed on split peas. Heredity *79*, 15–23.

Gressent, F., Rahioui, I., and Rahbe, Y. (2003). Characterization of a high-affinity binding site for the pea albumin 1b entomotoxin in the weevil Sitophilus. Eur. J. Biochem. *270*, 2429–2435.

Gressent, F., Duport, G., Rahioui, I., Pauchet, Y., Bolland, P., Specty, O., and Rahbe, Y. (2007). Biological Activity and Binding Site Characteristics of the PA1b Entomotoxin on Insects from Different Orders. J. Insect Sci. *7*, 1–10.

Gressent, F., Da Silva, P., Eyraud, V., Karaki, L., and Royer, C. (2011). Pea Albumin 1 Subunit b (PA1b), a Promising Bioinsecticide of Plant Origin. Toxins *3*, 1502–1517.

Gryspeirt, A., and Grégoire, J.-C. (2012). Effectiveness of the high dose/refuge strategy for managing pest resistance to Bacillus thuringiensis (Bt) plants expressing one or two toxins. Toxins *4*, 810–835.

Hagenmaier, H.-P., Werner, G., Drautz, H., Holst, H., Zahner, H., Brandes, W., Reinecke, P., Zoebelein, G., Stendel, W., Andrews, P., *et al*. (1985). Bafilomycin pesticides.

Hansen, A.K., and Moran, N.A. (2011). Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. Proc. Natl. Acad. Sci. *108*, 2849–2854.

Harju, S., Fedosyuk, H., and Peterson, K.R. (2004). Rapid isolation of yeast genomic DNA: Bust n' Grab. BMC Biotechnol. *4*, 8.

Hellmich, R.L., and Hellmich, K.A. (2012). Use and impact of Bt maize. Nat. Educ. Knowl. *3*, 4.

Hendriksma, H.P., Härtel, S., and Steffan-Dewenter, I. (2011). Testing Pollen of Single and Stacked Insect-Resistant Bt-Maize on In vitro Reared Honey Bee Larvae. PLoS ONE *6*, e28174.

Herzig, V., and King, G. (2015). The cystine knot is responsible for the exceptional stability of the insecticidal spider toxin ω-hexatoxin-Hv1a. Toxins 4366–4380.

Hester, G., Kaku, H., Goldstein, I.J., and Wright, C.S. (1995). Structure of mannose-specific snowdrop (Galanthus nivalis) lectin is representative of a new plant lectin family. Nat. Struct. Mol. Biol. *2*, 472–479.

Higgins, D.R., and Cregg, J.M. (1998). Introduction to Pichia pastoris. In Pichia Protocols, (New Jersey: Humana Press), pp. 1–16.

Higgins, T.J., Chandler, P.M., Randall, P.J., Spencer, D., Beach, L.R., Blagrove, R.J., Kortt, A.A., and Inglis, A.S. (1986). Gene structure, protein structure, and regulation of the synthesis of a sulfur-rich protein in pea seeds. J. Biol. Chem. *261*, 11124–11130.

Hilder, V.A., Powell, K.S., Gatehouse, A.M.R., Gatehouse, J.A., Gatehouse, L.N., Shi, Y., Hamilton, W.D.O., Merryweather, A., Newell, C.A., and Timans, J.C. (1995). Expression of snowdrop lectin in transgenic tobacco plants results in added protection against aphids. Transgenic Res. *4*, 18–25.

Hodge, S., and Powell, G. (2010). Conditional Facilitation of an Aphid Vector, *Acyrthosiphon pisum* , by the Plant Pathogen, *Pea Enation Mosaic Virus*. J. Insect Sci. *10*, 1–14.

Hogervorst, P.A.M., Ferry, N., Gatehouse, A.M.R., Wäckers, F.L., and Romeis, J. (2006). Direct effects of snowdrop lectin (GNA) on larvae of three aphid predators and fate of GNA after ingestion. J. Insect Physiol. *52*, 614–624.

Howarth, F.G. (1991). Environmental Impacts of Classical Biological Control. Annu. Rev. Entomol. *36*, 485–509.

Hu, W., Li, F., Yang, X., Li, Z., Xia, H., Li, G., Wang, Y., and Zhang, Z. (2004). A flexible peptide linker enhances the immunoreactivity of two copies HBsAg preS1 (21–47) fusion protein. J. Biotechnol. *107*, 83–90.

Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.S., Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., and Crea, R. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. Proc. Natl. Acad. Sci. *85*, 5879–5883.

Inan, M., and Meagher, M.M. (2001). Non-repressing carbon sources for alcohol oxidase (AOX1) promoter of Pichia pastoris. J. Biosci. Bioeng. *92*, 585–589.

Ippolito, A., Kattwinkel, M., Rasmussen, J.J., Schäfer, R.B., Fornaroli, R., and Liess, M. (2015). Modeling global distribution of agricultural insecticides in surface waters. Environ. Pollut. *198*, 54–60.

IRAC Insecticide resistance mechanisms.

ISAAA (2017). Global Status of Commercialized Biotech/GM Crops in 2017: Biotech Crop Adoption Surges as Economic Benefits Accumulate in 22 years (Ithaca).

Jahic, M., Gustavsson, M., Jansen, A.-K., Martinelle, M., and Enfors, S.-O. (2003). Analysis and control of proteolysis of a fusion protein in Pichia pastoris fed-batch processes. J. Biotechnol. *102*, 45–53.

James, C. (2010). brief 42 Global Status of Commercialized Biotech/GM Crops: 2010. 292.

Jolly, C.J., Shine, R., and Greenlees, M.J. (2015). The impact of invasive cane toads on native wildlife in southern Australia. Ecol. Evol. *5*, 3879–3894.

Jouvensal, L., Quillien, L., Ferrasson, E., Rahbé, Y., Guéguen, J., and Vovelle, F. (2003). PA1b, an Insecticidal Protein Extracted from Pea Seeds (*Pisum sativum*): [1] H-2-D NMR Study and Molecular Modeling [†,‡]. Biochemistry *42*, 11915–11923.

Jurat-Fuentes, J.L., and Adang, M.J. (2004). Characterization of a Cry1Ac-receptor alkaline phosphatase in susceptible and resistant Heliothis virescens larvae. Eur. J. Biochem. FEBS *271*, 3127–3135.

Kendall, H.W., and Pimentel, D. (1994). Constraints on the Expansion of the Global Food Supply. Ambio *23*, 198–205.

Khush, G.S. (2001). Green revolution: the way forward. Nat. Rev. Genet. *2*, 815–822.

Kimple, M.E., Brill, A.L., and Pasker, R.L. (2013). Overview of Affinity Tags for Protein Purification: Affinity Tags for Protein Purification. In Current Protocols in Protein Science, J.E. Coligan, B.M. Dunn, D.W. Speicher, and P.T. Wingfield, eds. (Hoboken, NJ, USA: John Wiley & Sons, Inc.), pp. 9.9.1-9.9.23.

Klein, U. (1992). THE INSECT V-ATPase, A PLASMA MEMBRANE PROTON PUMP ENERGIZING SECONDARY ACTIVE TRANSPORT: IMMUNOLOGICAL EVIDENCE FOR THE OCCURRENCE OF A V-ATPase IN INSECT ION-TRANSPORTING EPITHELIA. J. Exp. Biol. *172*, 345.

Knight, P.J., Crickmore, N., and Ellar, D.J. (1994). The receptor for Bacillus thuringiensis CrylA(c) delta-endotoxin in the brush border membrane of the lepidopteran Manduca sexta is aminopeptidase N. Mol. Microbiol. *11*, 429–436.

Koch, R.L. (2003). The multicolored Asian lady beetle, Harmonia axyridis: a review of its biology, uses in biological control, and non-target impacts. J. Insect Sci. Online *3*, 32.

Koganesawa, N., Aizawa, T., Shimojo, H., Miura, K., Ohnishi, A., Demura, M., Hayakawa, Y., Nitta, K., and Kawano, K. (2002). Expression and purification of a small cytokine growth-blocking peptide from armyworm Pseudaletia separata by an optimized fermentation method using the methylotrophic yeast Pichia pastoris. Protein Expr. Purif. *25*, 416–425.

Koureas, M., Tsakalof, A., Tsatsakis, A., and Hadjichristodoulou, C. (2012). Systematic review of biomonitoring studies to determine the association between exposure to organophosphorus and pyrethroid insecticides and human health outcomes. Toxicol. Lett. *210*, 155–168.

Kurtzman, C.P. (2005). Description of Komagataella phaffii sp. nov. and the transfer of Pichia pseudopastoris to the methylotrophic yeast genus Komagataella. Int. J. Syst. Evol. Microbiol. *55*, 973–976.

Larsson, H. (2005). A crop loss model and economic thresholds for the grain aphid, Sitobion avenae (F.), in winter wheat in southern Sweden. Crop Prot. *24*, 397–405.

Laus, M.C., Logman, T.J., Lamers, G.E., Van Brussel, A.A.N., Carlson, R.W., and Kijne, J.W. (2006). A novel polar surface polysaccharide from Rhizobium leguminosarum binds host plant lectin. Mol. Microbiol. *59*, 1704–1713.

Lawo, N.C., Wäckers, F.L., and Romeis, J. (2009). Indian Bt Cotton Varieties Do Not Affect the Performance of Cotton Aphids. PLoS ONE *4*, e4804.

Le Goff, G., and Giraudo, M. (2019). Effects of Pesticides on the Environment and Insecticide Resistance. In Olfactory Concepts of Insect Control - Alternative to Insecticides, J.-F. Picimbon, ed. (Cham: Springer International Publishing), pp. 51–78.

Legrand, A., and Barbosa, P. (2000). Pea Aphid (Homoptera: Aphididae) Fecundity, Rate of Increase, and Within-Plant Distribution Unaffected by Plant Morphology. Environ. Entomol. *29*, 987–993.

Lemaux, P.G. (2008). Genetically Engineered Plants and Foods: A Scientist's Analysis of the Issues (Part I). Annu. Rev. Plant Biol. *59*, 771–812.

van Lenteren, J.C. (2012). The state of commercial augmentative biological control: plenty of natural enemies, but a frustrating lack of uptake. BioControl *57*, 1–20.

van Lenteren, J.C., and Bueno, V.H.P. (2003). Augmentative biological control of arthropods in Latin America. BioControl *48*, 123–139.

van Lenteren, J.C., van Roermund, H.J.W., and Sütterlin, S. (1996). Biological Control of Greenhouse Whitefly (Trialeurodes vaporariorum) with the ParasitoidEncarsia formosa:How Does It Work? Biol. Control *6*, 1–10.

Li, H., Chougule, N.P., and Bonning, B.C. (2011). Interaction of the Bacillus thuringiensis delta endotoxins Cry1Ac and Cry3Aa with the gut of the pea aphid, Acyrthosiphon pisum (Harris). J. Invertebr. Pathol. *107*, 69–78.

Li, P., Anumanthan, A., Gao, X.-G., Ilangovan, K., Suzara, V.V., Düzgüneş, N., and Renugopalakrishnan, V. (2007). Expression of Recombinant Proteins in Pichia Pastoris. Appl. Biochem. Biotechnol. *142*, 105–124.

Li, Z., Du, L., Zhang, W., Zhang, X., Jiang, Y., Liu, K., Men, P., Xu, H., Fortman, J.L., Sherman, D.H., *et al*. (2017). Complete elucidation of the late steps of bafilomycin biosynthesis in *Streptomyces lohii*. J. Biol. Chem. *292*, 7095–7104.

Lichty, J.J., Malecki, J.L., Agnew, H.D., Michelson-Horowitz, D.J., and Tan, S. (2005). Comparison of affinity tags for protein purification. Protein Expr. Purif. *41*, 98–105.

Liu, N., and Yue, X. (2000). Insecticide Resistance and Cross-Resistance in the House Fly (Diptera: Muscidae). J. Econ. Entomol. *93*, 1269–1275.

de Lourdes Corrêa Figueiredo, M., Cruz, I., da Silva, R.B., and Foster, J.E. (2015). Biological control with Trichogramma pretiosum increases organic maize productivity by 19.4%. Agron. Sustain. Dev. *35*, 1175–1183.

Lu, Y.H., Wu, K.M., Wyckhuys, K.A.G., and Guo, Y.Y. (2009). Potential of mungbean, Vigna radiatus as a trap crop for managing Apolygus lucorum (Hemiptera: Miridae) on Bt cotton. Crop Prot. *28*, 77–81.

de Maagd, R. (2001). How Bacillus thuringiensis has evolved specific toxins to colonize the insect world. Trends Genet. *17*, 193–199.

Macauley-Patrick, S., Fazenda, M.L., McNeil, B., and Harvey, L.M. (2005). Heterologous protein production using thePichia pastoris expression system. Yeast *22*, 249–270.

Manandhar, R., and Wright, M.G. (2016). Effects of Interplanting Flowering Plants on the Biological Control of Corn Earworm (Lepidoptera: Noctuidae) and Thrips (Thysanoptera: Thripidae) in Sweet Corn. J. Econ. Entomol. *109*, 113–119.

Mariyono, J. (2008). Direct and indirect impacts of integrated pest management on pesticide use: a case of rice agriculture in Java, Indonesia. Pest Manag. Sci. *64*, 1069–1073.

Marshansky, V., and Futai, M. (2008). The V-type H+-ATPase in vesicular trafficking: targeting, regulation and function. Curr. Opin. Cell Biol. *20*, 415–426.

Martinez-Torres, D., Foster, S.P., Field, L.M., Devonshire, A.L., and Williamson, M.S. (1999). A sodium channel point mutation is associated with resistance to DDT and pyrethroid insecticides in the peach-potato aphid, Myzus persicae (Sulzer) (Hemiptera: Aphididae). Insect Mol. Biol. *8*, 339–346.

Mathers, T.C., Chen, Y., Kaithakottil, G., Legeai, F., Mugford, S.T., Baa-Puyoulet, P., Bretaudeau, A., Clavijo, B., Colella, S., Collin, O., *et al*. (2017). Rapid transcriptional plasticity of duplicated gene clusters enables a clonally reproducing aphid to colonise diverse plant species. Genome Biol. *18*.

Mattanovich, D., Graf, A., Stadlmann, J., Dragosits, M., Redl, A., Maurer, M., Kleinheinz, M., Sauer, M., Altmann, F., and Gasser, B. (2009). Genome, secretome and glucose transport highlight unique features of the protein production host Pichia pastoris. Microb. Cell Factories *8*, 29.

McVean, R.I.K., and Dixon, A.F.G. (2002). The host plant range of the pea aphid subspecies Acyrthosiphon pisum ssp. destructor (Johnson) (Hom., Aphididae). J. Appl. Entomol. *126*, 281–286.

Mehrparvar, M., Zytynska, S.E., and Weisser, W.W. (2013). Multiple Cues for Winged Morph Production in an Aphid Metacommunity. PLoS ONE *8*, e58323.

Meier, T. (2005). Structure of the Rotor Ring of F-Type Na+-ATPase from Ilyobacter tartaricus. Science *308*, 659–662.

Men, X., Ge, F., Edwards, C.A., and Yardim, E.N. (2005). The influence of pesticide applications on Helicoverpa armigera Hübner and sucking pests in transgenic Bt cotton and non-transgenic cotton in China. Crop Prot. *24*, 319–324.

Miles, P.W. (1999). Aphid saliva. Biol. Rev. *74*, 41–85.

Mizell III, R.F. (2007). IMPACT OF HARMONIA AXYRIDIS (COLEOPTERA: COCCINELLIDAE) ON NATIVE ARTHROPOD PREDATORS IN PECAN AND CRAPE MYRTLE. Fla. Entomol. *90*, 524–536.

Mohanty, A.K., and Wiener, M.C. (2004). Membrane protein expression and production: effects of polyhistidine tag length and position. Protein Expr. Purif. *33*, 311–325.

Morton, R.L., Whitecross, M., and Higgins, T.J.V. (1993). Post-transcriptional control of the expression of a plant gene by an environmental factor: Sulphur regulation of the expression of the Pea Albumin 1 gene. In Genetic Aspects of Plant Mineral Nutrition, P.J. Randall, E. Delhaize, R.A. Richards, and R. Munns, eds. (Dordrecht: Springer Netherlands), pp. 311–319.

Muench, S.P., Huss, M., Song, C.F., Phillips, C., Wieczorek, H., Trinick, J., and Harrison, M.A. (2009). Cryo-electron Microscopy of the Vacuolar ATPase Motor Reveals its Mechanical and Regulatory Complexity. J. Mol. Biol. *386*, 989–999.

Muench, S.P., Rawson, S., Eyraud, V., Delmas, A.F., Da Silva, P., Phillips, C., Trinick, J., Harrison, M.A., Gressent, F., and Huss, M. (2014). PA1b Inhibitor Binding to Subunits *c* and *e* of the Vacuolar ATPase Reveals Its Insecticidal Mechanism. J. Biol. Chem. *289*, 16399–16408.

Murata, T. (2005). Structure of the Rotor of the V-Type Na+-ATPase from Enterococcus hirae. Science *308*, 654–659.

Nakasu, E.Y.T., Williamson, S.M., Edwards, M.G., Fitches, E.C., Gatehouse, J.A., Wright, G.A., and Gatehouse, A.M.R. (2014a). Novel biopesticide based on a spider venom peptide shows no adverse effects on honeybees. Proc. Biol. Sci. *281*.

Nakasu, E.Y.T., Edwards, M.G., Fitches, E.C., Gatehouse, J.A., and Gatehouse, A.M.R. (2014b). Transgenic plants expressing Ï‰-ACTX-Hv1a and snowdrop lectin (GNA) fusion protein show enhanced resistance to aphids. Front. Plant Sci. *5*.

Nault, L.R. (1997). Arthropod Transmission of Plant Viruses: a New Synthesis. Ann. Entomol. Soc. Am. *90*, 521–541.

Ndakidemi, B., Mtei, K., and Ndakidemi, P.A. (2016). Impacts of Synthetic and Botanical Pesticides on Beneficial Insects. Agric. Sci. *07*, 364–372.

Nelson, D.L. (1976). EVALUATION OF RAPESEED FLOUR AND PEA PROTEIN CONCENTRATE AS PROTEIN SUPPLEMENTS FOR HONEY BEES (HYMENOPTERA: APIDAE). Can. Entomol. *108*, 845–848.

Nelson, N., Perzov, N., Cohen, A., Hagai, K., Padler, V., and Nelson, H. (2000). The Cellular Biology of Proton-Motive Force Generation By V-ATPases. J. Exp. Biol. *203*, 89–95.

NFU (2017). IPM Plan.

Nishi, T., and Forgac, M. (2002). The vacuolar (H+)-ATPases — nature's most versatile proton pumps. Nat. Rev. Mol. Cell Biol. *3*, 94–103.

Norton, R.S., and Pallaghy, P.K. (1998). The cystine knot structure of ion channel toxins and related polypeptides. Toxicon Off. J. Int. Soc. Toxinology *36*, 1573–1583.

O'Donnell, M. (2017). The V-ATPase in insect epithelia. J. Exp. Biol. *220*, 3201–3203.

Oerke, E.C. (2005). Crop losses to pests. J. Agric. Sci. *144*, 31–43.

Oerke, E.C., and Dehne, H.W. (2004). Safeguarding production—losses in major crops and the role of crop protection. Crop Prot. *23*, 275–285.

Oliveira, A.C. de, Vidal, B. de C., and Sgarbieri, V.C. (1989). Lesions of intestinal epithelium by ingestion of bean lectins in rats. J. Nutr. Sci. Vitaminol. (Tokyo) *35*, 315–322.

Pang, R., Chen, M., Liang, Z., Yue, X., Ge, H., and Zhang, W. (2016). Functional analysis of CYP6ER1, a P450 gene associated with imidacloprid resistance in Nilaparvata lugens. Sci. Rep. *6*.

Pence, R.J. (1965). The Antimetabolite: Imidazole as a pesticide. Calif. Agric. *19*, 13–15.

Peshin, R., and Zhang, W. (2014). Integrated Pest Management and Pesticide Use. In Integrated Pest Management, D. Pimentel, and R. Peshin, eds. (Dordrecht: Springer Netherlands), pp. 1–46.

Peumans, W.J., and Van Damme, E.J. (1995). Lectins as plant defense proteins. Plant Physiol. *109*, 347–352.

Phaff, H.J., Miller, M.W., and Shifrine, M. (1956). The taxonomy of yeasts isolated fromDrosophila in the Yosemite region of California. Antonie Van Leeuwenhoek *22*, 145–161.

Pickering, G.J., Lin, Y., Reynolds, A., Soleas, G., Riesen, R., and Brindle, I. (2005). The Influence of *Harmonia axyridis* on Wine Composition and Aging. J. Food Sci. *70*, S128–S135.

Pigott, C.R., and Ellar, D.J. (2007). Role of Receptors in Bacillus thuringiensis Crystal Toxin Activity. Microbiol. Mol. Biol. Rev. *71*, 255.

Pilkington, L.J., Messelink, G., van Lenteren, J.C., and Le Mottee, K. (2010). "Protected Biological Control" – Biological pest management in the greenhouse industry. Biol. Control *52*, 216–220.

Ponsen, M.B. (1972). The site of potato leafroll virus multiplication in its vector, Myzus persicae : an anatomical study. Veenman.

Porcar, M., Grenier, A.-M., Federici, B., and Rahbe, Y. (2009). Effects of Bacillus thuringiensis - Endotoxins on the Pea Aphid (Acyrthosiphon pisum). Appl. Environ. Microbiol. *75*, 4897–4900.

Potgieter, T.I., Cukan, M., Drummond, J.E., Houston-Cummings, N.R., Jiang, Y., Li, F., Lynaugh, H., Mallem, M., McKelvey, T.W., Mitchell, T., *et al*. (2009). Production of monoclonal antibodies by glycoengineered Pichia pastoris. J. Biotechnol. *139*, 318–325.

Poulsen, M., Kroghsbo, S., Schrøder, M., Wilcks, A., Jacobsen, H., Miller, A., Frenzel, T., Danier, J., Rychlik, M., Shu, Q., *et al*. (2007). A 90-day safety study in Wistar rats fed genetically modified rice expressing snowdrop lectin Galanthus nivalis (GNA). Food Chem. Toxicol. *45*, 350–363.

Powell, K.S., Gatehouse, A.M.R., Hilder, V.A., and Gatehouse, J.A. (1993). Antimetabolic effects of plant lectins and plant and fungal enzymes on the nymphal stages of two important rice pests, Nilaparvata lugens and Nephotettix cinciteps. Entomol. Exp. Appl. *66*, 119–126.

Powell, K.S., Gatehouse, A.M.R., Hilder, V.A., and Gatehouse, J.A. (1995). Antifeedant effects of plant lectins and an enzyme on the adult stage of the rice brown planthopper, Nilaparvata lugens. Entomol. Exp. Appl. *75*, 51–59.

Powell, K.S., Spence, J., Bharathi, M., Gatehouse, J.A., and Gatehouse, A.M.R. (1998). Immunohistochemical and developmental studies to elucidate the mechanism of action of the snowdrop lectin on the rice brown planthopper, *Nilaparvata lugens* (Stal). J. Insect Physiol. *44*, 529–539.

Powell, M.E., Bradish, H.M., Cao, M., Makinson, R., Brown, A.P., Gatehouse, J.A., and Fitches, E.C. (2019). Demonstrating the potential of a novel spider venom-based biopesticide for target-specific control of the small hive beetle, a serious pest of the European honeybee. J. Pest Sci.

Prielhofer, R., Cartwright, S.P., Graf, A.B., Valli, M., Bill, R.M., Mattanovich, D., and Gasser, B. (2015). Pichia pastoris regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. BMC Genomics *16*.

Promdonkoy, B., and Ellar, D.J. (2003). Investigation of the pore-forming mechanism of a cytolytic δ-endotoxin from Bacillus thuringiensis. Biochem. J. *374*, 255–259.

Puinean, A.M., Foster, S.P., Oliphant, L., Denholm, I., Field, L.M., Millar, N.S., Williamson, M.S., and Bass, C. (2010). Amplification of a Cytochrome P450 Gene Is Associated with Resistance to Neonicotinoid Insecticides in the Aphid Myzus persicae. PLoS Genet. *6*, e1000999.

Purcell, J.P., and Perlack, F.J. (2004). Global impact of insect-resistant (bt) cotton. J. Agrobiotechnology Manag. Econ. *7*, 27–30.

Pyati, P., Fitches, E.C., and Gatehouse, J.A. (2014). Optimising expression of the recombinant fusion protein biopesticide ω-hexatoxin-Hv1a/GNA in *Pichia pastoris*: sequence modifications and a simple method for the generation of multi-copy strains. J. Ind. Microbiol. Biotechnol. *41*, 1237–1247.

Qaim, M., and Zilberman, D. (2003). Yeild effects of genetically modified crops in developing countries. Science *299*, 900–902.

Qiao, F., and Yao, Y. (2015). Is the economic benefit of Bt cotton dying away in China. China Agric. Econ. Rev. *7*, 322–336.

Raemaekers, R.J.M., de Muro, L., Gatehouse, J.A., and Fordham-Skelton, A.P. (1999). Functional phytohemagglutinin (PHA) and *Galanthus nivalis* agglutinin (GNA) expressed in *Pichia pastoris*. Correct N-terminal processing and secretion of heterologous proteins expressed using the PHA-E signal peptide. Eur. J. Biochem. *265*, 394–403.

Rahbé, Y., Sauvion, N., Febvay, G., Peumans, W.J., and Gatehouse, A.M.R. (1995a). Toxicity of lectins and processing of ingested proteins in the pea aphid Acyrthosiphon pisum. Entomol. Exp. Appl. *76*, 143–155.

Rahbé, Y., Sauvion, N., Febvay, G., Peumans, W.J., and Gatehouse, A.M.R. (1995b). Toxicity of lectins and processing of ingested proteins in the pea aphid Acyrthosiphon pisum. Entomol. Exp. Appl. *76*, 143–155.

Rahbé, Y., Ferrasson, E., Rabesona, H., and Quillien, L. (2003). Toxicity to the pea aphid Acyrthosiphon pisum of anti-chymotrypsin isoforms and fragments of Bowman–Birk protease inhibitors from pea seeds. Insect Biochem. Mol. Biol. *33*, 299–306.

Rahioui, I., Eyraud, V., Karaki, L., Sasse, F., Carre-Pierrat, M., Qin, A., Zheng, M.H., Toepfer, S., Sivignon, C., Royer, C., *et al*. (2014). Host range of the potential biopesticide Pea Albumin 1b (PA1b) is limited to insects. Toxicon *89*, 67–76.

Ramsey, J.S., Rider, D.S., Walsh, T.K., De Vos, M., Gordon, K.H.J., Ponnala, L., Macmil, S.L., Roe, B.A., and Jander, G. (2010). Comparative analysis of detoxification enzymes in Acyrthosiphon pisum and Myzus persicae: Aphid detoxification enzymes. Insect Mol. Biol. *19*, 155–164.

Rao, K.V., Rathore, K.S., Hodges, T.K., Fu, X., Stoger, E., Sudhakar, D., Williams, S., Christou, P., Bharathi, M., Bown, D.P., *et al*. (1998). Expression of snowdrop lectin (GNA) in transgenic rice plants confers resistance to rice brown planthopper. Plant J. Cell Mol. Biol. *15*, 469–477.

Rausch, M.A., Chougule, N.P., Deist, B.R., and Bonning, B.C. (2016). Modification of Cry4Aa toward Improved Toxin Processing in the Gut of the Pea Aphid, Acyrthosiphon pisum. PLOS ONE *11*, e0155466.

Rebnegger, C., Graf, A.B., Valli, M., Steiger, M.G., Gasser, B., Maurer, M., and Mattanovich, D. (2014). In *Pichia pastoris* , growth rate regulates protein synthesis and secretion, mating and stress response. Biotechnol. J. *9*, 511–525.

Rodrigues, D., Pillaca-Pullo, O., Torres-Obreque, K., Flores-Santos, J., Sánchez-Moguel, I., Pimenta, M.V., Basi, T., Converti, A., Lopes, A.M., Monteiro, G., *et al*. (2019). Fed-Batch Production of Saccharomyces cerevisiae L-Asparaginase II by Recombinant Pichia pastoris MUTs Strain. Front. Bioeng. Biotechnol. *7*.

Roh, J.Y., Choi, J.Y., Li, M.S., Jin, B.R., and Je, Y.H. (2007). Bacillus thuringiensis as a specific, safe, and effective tool for insect pest control. J. Microbiol. Biotechnol. *17*, 547–559.

Royer, C., Foy, S., Da Silva, P., Gressent, F., Karaki, L., and Rahbé, Y. (2018). ENTOMOTOXIC POLYPEPTIDES.

Sabarwal, A., Kumar, K., and Singh, R.P. (2018). Hazardous effects of chemical pesticides on human health–Cancer and other associated disorders. Environ. Toxicol. Pharmacol. *63*, 103–114.

Sadeghi, A., Van Damme, E.J., Michiels, K., Kabera, A., and Smagghe, G. (2009). Acute and chronic insecticidal activity of a new mannose-binding lectin from *Allium porrum* against *Acyrthosiphon pisum via* an artificial diet. Can. Entomol. *141*, 95–101.

Sanahuja, G., Banakar, R., Twyman, R.M., Capell, T., and Christou, P. (2011). Bacillus thuringiensis: a century of research, development and commercial applications: A century of Bacillus thuringiensis. Plant Biotechnol. J. *9*, 283–300.

Sauvion, N., Rahbé, Y., Peumans, W.J., Van Damme, E.J., Gatehouse, J.A., and Gatehouse, A.M.R. (1996). Effects of GNA and other mannose binding lectins on development and fecundity of the peach-potato aphid Myzus persicae. Entomol. Exp. Appl. *79*, 285–293.

Sauvion, N., Nardon, C., Febvay, G., Gatehouse, A.M.R., and Rahbé, Y. (2004). Binding of the insecticidal lectin Concanavalin A in pea aphid, Acyrthosiphon pisum (Harris) and induced effects on the structure of midgut epithelial cells. J. Insect Physiol. *50*, 1137–1150.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., *et al*. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.

Schwarz, F., and Aebi, M. (2011). Mechanisms and principles of N-linked protein glycosylation. Curr. Opin. Struct. Biol. *21*, 576–582.

Schwinghamer, M.W., Nicholas, A.H., and Schilg, M.A. (2009). Three aphid vectors of faba bean (*Vicia faba*) viruses in northern New South Wales and occurrence of *Acyrthosiphon pisum* -transmitted isolates of *Soybean dwarf virus*. Australas. Plant Pathol. *38*, 262.

Simon, J.C., Dedryver, C.A., Pierre, J.S., Tanguy, S., and Wegorek, P. (1991). The influence of clone and morph on the parameters of intrinsic rate of increase in the cereal aphids Sitobion avenae and Rhopalosiphum padi. Entomol. Exp. Appl. *58*, 211–220.

Slater, R., Paul, V.L., Andrews, M., Garbay, M., and Camblin, P. (2011). Identifying the presence of neonicotinoidresistant peach-potato aphid (Myzus persicae) in the peach-growing regions of southern France and northern Spain. Pest Manag. Sci. *68*, 634–638.

Smeekens, J.M., Xiao, H., and Wu, R. (2017). Global Analysis of Secreted Proteins and Glycoproteins in *Saccharomyces cerevisiae*. J. Proteome Res. *16*, 1039–1049.

Smith, S.M. (1996). Biological Control with Trichogramma: Advances, Successes, and Potential of Their Use. Annu. Rev. Entomol. *41*, 375–406.

Stoger, E., Williams, S., Christou, P., Down, R.E., and Gatehouse, J.A. (1999). Expression of the insecticidal lectin from snowdrop (Galanthus nivalis agglutinin; GNA) in transgenic wheat plants: effects on predation by the grain aphid Sitobion avenae. Mol. Breed. *5*, 65–73.

Strycharz, J.P., Lao, A., Li, H., Qiu, X., Lee, S.H., Sun, W., Yoon, K.S., Doherty, J.J., Pittendrigh, B.R., and Clark, J.M. (2013). Resistance in the highly DDT-resistant 91-R strain of Drosophila melanogaster involves decreased penetration, increased metabolism, and direct excretion. Pestic. Biochem. Physiol. *107*, 207–217.

Sumi, A., Okuyama, K., Kobayashi, K., Ohtani, W., Ohmura, T., and Yokoyama, K. (1999). Purification of recombinant human serum albumin efficient purification using STREAMLINE. Bioseparation *8*, 195–200.

Tabashnik, B.E. (1994). Evolution of Resistance to Bacillus Thuringiensis. Annu. Rev. Entomol. *39*, 47–79.

Tabashnik, B.E., and Carrière, Y. (2019). Global Patterns of Resistance to Bt Crops Highlighting Pink Bollworm in the United States, China, and India. J. Econ. Entomol.

Tabashnik, B.E., Liu, Y.-B., Finson, N., Masson, L., and Heckel, D.G. (1997). One gene in diamondback moth confers resistance to four Bacillus thuringiensis toxins. Proc. Natl. Acad. Sci. *94*, 1640–1644.

Tedford, H.W., Sollod, B.L., Maggio, F., and King, G.F. (2004). Australian funnel-web spiders: master insecticide chemists. Toxicon Off. J. Int. Soc. Toxinology *43*, 601–618.

Thomas, W.E., and Ellar, D.J. (1983). Mechanism of action of Bacillus thuringiensis var israelensis insecticidal δ-endotoxin. FEBS Lett. *154*, 362–368.

Tilman, D., Blazer, C., Hill, J., and Befort, B. (2011). Global food demand and the sustainable intensification of agriculture. Proc. Natl. Acad. Sci. 20260–20264.

Tjallingii, W.F. (2006). Salivary secretions by aphids interacting with proteins of phloem wound responses. J. Exp. Bot. *57*, 739–745.

Tomlinson, J.A. (1987). Epidemiology and control of virus diseases of vegetables. Ann. Appl. Biol. *110*, 661–681.

Trung, N., Fitches, E.C., and Gatehouse, J.A. (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 Biotechnol. *6*, 18.

Tschopp, J.F., Brust, P.F., Cregg, J.M., Stillman, C.A., and Gingeras, T.R. (1987). Expression of the lacZ gene from two methanol-regulated promoters in Pichia pastoris. Nucleic Acids Res. *15*, 3859–3876.

United Nations, D. of E. and S.A., Population division (2019). World Population Prospects 2019: Data Booklet (UN).

Upadhyay, S.K., Chandrashekar, K., Thakur, N., Verma, P.C., Borgio, J.F., Singh, P.K., and Tuli, R. (2011). RNA interference for the control of whiteflies (Bemisia tabaci) by oral route. J. Biosci. *36*, 153–161.

US EPA, O. (2014). Introduction to Integrated Pest Management.

Vadlamudi, R.K., Weber, E., Ji, I., Ji, T.H., and Bulla, L.A. (1995). Cloning and expression of a receptor for an insecticidal toxin of Bacillus thuringiensis. J. Biol. Chem. *270*, 5490–5494.

Van Damme, E.J., Allen, A.K., and Peumans, W.J. (1987). Isolation and characterization of a lectin with exclusive specificity towards mannose from snowdrop (Galanthus nivalis) bulbs. FEBS Lett. *215*, 140–144.

Van Damme, E.J., Peumans, W.J., Barre, A., and Rougé, P. (1998). Plant lectins: A composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles. Crit. Rev. Plant Sci. *17*, 575–692.

Vanz, A.L., Lünsdorf, H., Adnan, A., Nimtz, M., Gurramkonda, C., Khanna, N., and Rinas, U. (2012). Physiological response of Pichia pastoris GS115 to methanol-induced high level production of the Hepatitis B surface antigen: catabolic adaptation, stress responses, and autophagic processes. Microb. Cell Factories *11*, 103.

Wakefield, M.E., Fitches, E.C., Bell, H.A., and Gatehouse, A.M.R. (2010). The snowdrop lectin Galanthus nivalis agglutinin (GNA) and a fusion protein ButaIT/GNA have a differential affect on a pest noctuid Lacanobia oleracea and the ectoparasitoid Eulophus pennicornis. Physiol. Entomol. *35*, 334–342.

Waterham, H.R., Digan, M.E., Koutz, P.J., Lair, S.V., and Cregg, J.M. (1997). Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene *186*, 37–44.

Watt, A.D. (1979). The effect of cereal growth stages on the reproductive activity of Sitobion avenue and Metopolophium dirhodum. Ann. Appl. Biol. *91*, 147–157.

Weinacker, D., Rabert, C., Zepeda, A.B., Figueroa, C.A., Pessoa, A., and Farías, J.G. (2014). Applications of recombinant Pichia pastoris in the healthcare industry. Braz. J. Microbiol. Publ. Braz. Soc. Microbiol. *44*, 1043–1048.

Werten, M.W.T., van den Bosch, T.J., Wind, R.D., Mooibroek, H., and de Wolf, F.A. (1999). High-yield secretion of recombinant gelatins byPichia pastoris. Yeast *15*, 1087–1096.

Whalon, M.E., Mota-Sanchez, D., and Hollingworth, R.M. (2008). Global Pesticide Resistance in Arthropods (CABI).

Whitfield, A.E., Falk, B.W., and Rotenberg, D. (2015). Insect vector-mediated transmission of plant viruses. Virology *479–480*, 278–289.

Wieczorek, H., Grber, G., Harvey, W.R., Huss, M., Merzendorfer, H., and Zeiske, W. (2000). Structure and regulation of insect plasma membrane H(+)V-ATPase. J. Exp. Biol. *203*, 127–135.

Wieczorek, H., Beyenbach, K.W., Huss, M., and Vitavska, O. (2009). Vacuolar-type proton pumps in insect epithelia. J. Exp. Biol. *212*, 1611–1619.

Will, T., Tjallingii, W.F., Thonnessen, A., and van Bel, A.J.E. (2007). Molecular sabotage of plant defense by aphid saliva. Proc. Natl. Acad. Sci. *104*, 10536–10541.

Williams, I.S., Dewar, A.M., Dixon, A.F.G., and Thornhill, W.A. (2000). Alate production by aphids on sugar beet: how likely is the evolution of sugar beet-specific biotypes? J. Appl. Ecol. *37*, 40–51.

Wilson, A.C.C., Ashton, P.D., Calevro, F., Charles, H., Colella, S., Febvay, G., Jander, G., Kushlan, P.F., Macdonald, S.J., Schwartz, J.F., *et al*. (2010). Genomic insight into the amino acid relations of the pea aphid, Acyrthosiphon pisum, with its symbiotic bacterium Buchnera aphidicola: Amino acid metabolism genes in the pea aphid symbiosis. Insect Mol. Biol. *19*, 249–258.

Woodcock, B.A., Bullock, J.M., Shore, R.F., Heard, M.S., Pereira, M.G., Redhead, J., Ridding, L., Dean, H., Sleep, D., Henrys, P., *et al*. (2017). Country-specific effects of neonicotinoid pesticides on honey bees and wild bees. Science *356*, 1393–1395.

Xu, L., Wang, Z., Zhang, J., He, K., Ferry, N., and Gatehouse, A.M.R. (2010). Cross-resistance of Cry1Ab-selected Asian corn borer to other Cry toxins: Cross-resistance to Bt toxins in Ostrinia furnacalis. J. Appl. Entomol. *134*, 429–438.

Yamada, Y., Matsuda, M., Maeda, K., and Mikata, K. (1995). The Phylogenetic Relationships of Methanol-assimilating Yeasts Based on the Partial Sequences of 18S and 26S Ribosomal RNAs: The Proposal of *Komagataella* Gen. Nov. (Saccharomycetaceae). Biosci. Biotechnol. Biochem. *59*, 439–444.

Yamazaki, T., Takaoka, M., Katoh, E., Hanada, K., Sakita, M., Sakata, K., Nishiuchi, Y., and Hirano, H. (2003). A possible physiological function and the tertiary structure of a 4-kDa peptide in legumes. Eur. J. Biochem. *270*, 1269–1276.

Yang, S. (2015). Exploitation of small cysteine-rich spider protein toxins as bio-insecticides. Doctoral Thesis. Durham University.

Yang, S., Pyati, P., Fitches, E.C., and Gatehouse, J.A. (2014). A recombinant fusion protein containing a spider toxin specific for the insect voltage-gated sodium ion channel shows oral toxicity towards insects of different orders. Insect Biochem. Mol. Biol. *47*, 1–11.

Zahrl, R.J., Peña, D.A., Mattanovich, D., and Gasser, B. (2017). Systems biotechnology for protein production in Pichia pastoris. FEMS Yeast Res. *17*.

Zalucki, M.P., Adamson, D., and Furlong, M.J. (2009). The future of IPM: whither or wither? Aust. J. Entomol. *48*, 85–96.

Zapata, N., Van Damme, E.J., Vargas, M., Devotto, L., and Smagghe, G. (2016). Insecticidal activity of a protein extracted from bulbs of Phycella australis Ravenna against the aphids Acyrthosiphon pisum Harris and Myzus persicae Sulzer. Chil. J. Agric. Res. *76*, 188–194.

Zhao, J.-Z., Cao, J., Collins, H.L., Bates, S.L., Roush, R.T., Earle, E.D., and Shelton, A.M. (2005). Concurrent use of transgenic plants expressing a single and two Bacillus thuringiensis genes speeds insect adaptation to pyramided plants. PNAS *102*, 8426–8430.

Zhu, F., Lavine, L., O'Neal, S., Lavine, M., Foss, C., and Walsh, D. (2016). Insecticide Resistance and Management Strategies in Urban Ecosystems. Insects *7*, 2.

Arthropod Pesticide Resistance Database | Michigan State University.