Investigating the specificity of *Bacillus thuringiensis* Cry41Aa toxin through mutagenesis in domain II loops

Alicia Elhigazi

Submitted for the award of Degree of Doctor of Philosophy

Department of Biochemistry

School of Life Sciences

University of Sussex

February 2019

Work not submitted elsewhere for examination

I hereby declare that this thesis has not been submitted in whole or in part to this or any

other University for the award of a degree

Alicia Elhigazi

Dedication

This work is dedicated to my loving father Higazi Ali Elhigazi, whom sparked science in me at a young age. He was an excellent scientist, and a wonderful human who left this world too soon. I pray this makes him proud.

Acknowledgements

I am thankful to my supervisor, Dr Neil Crickmore, for his patience, guidance and support. His teachings and remarkably consideration were essential in the completion of this work. I would like to extend my gratitude to my wonderful lab sister Barbara Domanska for her help with cell assay and electrophysiology work, and so much more. Her continued support and kindness made this experience wonderful part of my life. I thank Eva Fortea and Prof Jean-Louis Schwartz, university of Montreal, Canada.

 I am thankful to my lab colleagues Vidisha Krishnan, Heather Collins, Aminah Barqawi, and Wided Souissi. Their input and discussion were always educational and helpful. I would like to extend my gratitude to Michelle West for cell assay work, Julian Thorpe for transmission electron microscopy work and Simon Morley for donation of antibody. I thank Dr Taseen S. Design and Dr Hafejee Essackjee for thoguful feedback and being great neighbours.

For their unwavering belief and love I am eternally grateful to my husband Sajjad Alhawsawi and my mother Madeleine Castrillon. Their support and patience were key to the completion of this work. To my lovely brother Amin Elhegazi and fabulous children, I am grateful for their love, patience and understanding during my PhD.

I thank God for this journey, the learning that came from it, and the wonderful people who walked along my path.

Abstract

Bacillus thuringiensis produces a range of toxins that include both the insecticidal Cry toxins, non-insecticidal, non-haemolytic, and Parasporins. The latter exhibits cytocidal activity to some cancer cell lines. Parasporins 3 or Cry41Aa is cytocidal to human hepatic HepG2 cell lines. It contains the five conserved sequence blocks found in many insecticidal 3-domain Cry toxins and is also believed to possess the same 3-domain structure. In addition, it has an extra loop in its domain II as well as an additional ricin domain at its C-terminus. studies on insecticidal Cry toxins have implicated domain II loops in the specificity of a toxin to target a particular cell. In this study the specificity of Cry41Aa towards HepG2 cell lines was investigated. Bioinformatic tools were used to predict domain II loops of Cry41Aa. A number of mutants were created to investigate its specificity. Loop exchange mutants between loop 3 Cry41Aa and Cry loop 3 of insecticidal were created but did not result in a proteolytically stable protein. Domain II hybrids of Cry41Aa and insecticidal Cry toxins were created but these did not result in a proteolytic stable protein. Finally, residue substitutions with alanine in loop 1,3, and the extra loop resulted in stable activated toxins. loop 1 mutants retained toxicity. The extra loop mutant lost toxicity towards HepG2 cell lines. A number of Loop 3 mutants were made. Recombinant, Y514A and W511F retained toxicity towards HepG2 cells. Recombinant W511A and several recombinants at position 509 including F509A did not exert toxicity as confirmed by cell viability assays. Despite the lack of toxicity, membrane damage assays and western blots on HepG2 incubated with F509A revealed the likely presence of pores and phosphorylation of p38. Cell electrophysiology tools were applied to investigate the effect that nontoxic recombinant F509A on artificial and cell

membrane. Cry41Aa induced the formation of stable pores, cell membrane damage and subsequent cell death. F509A induced the formation of unstable pores and did not compromise the integrity of cell membrane. The study Findings indicated that Cry41Aa is likely to have a similar mode of action as insecticidal Cry toxins.

List of abbreviations

A – ampere

BBMV - brush border membrane vesicle

bp - base pairs

BSA - bovine serum albumin

Bt – Bacillus thuringiensis

°C - degrees Celsius

cAMP - 3', 5' - cyclic adenosine monophosphate

CAPS – 3 - (cyclohexylamino) – 1 - propanesulfonic acid

Cry - crystal

Cyt – cytolytic

Da - Dalton

DMEM - Dulbecco's modified Eagle medium

DMSO - dimethyl sulfoxide

DNA - deoxyribonucleic acid

DPBS - Dulbecco's phosphate-buffered saline

DTT - dithiothreitol

EC_{50} - half maximal effective concentration

ECL - enhanced chemiluminescence

E.coli - Escherichia coli

EDTA – 2, 2', 2'', 2''' - (Ethane - 1, 2 - diyldinitrilo) tetraacetic acid

EGTA – ethylene glycol – bis (2 - aminoethylether) - *N,N,N',N'* - tetraacetic acid

F - farad

FCS – fetal calf serum

FPLC - fast protein liquid chromatography

G - conductance

HEPES – 2 - [4 - (2 - hydroxyethyl) piperazin – 1 - yl] ethanesulfonic acid

HRP - horseradish peroxidase

MEM – modified Eagle medium

NP-40 – nonidet-P40

OD_{600} - optical density measured at 600 nm

PAGE - polyacrylamide gel electrophoresis

PBS – phosphate - buffered saline

PBS-T – phosphate - buffered saline with Tween - 20

PC - phosphatidylcholine

PCR – polymerase chain reaction

PE - phosphatidylethanolamine

PEG - polyethylene glycol

pI - isoelectric point

PLB – planar lipid bilayer

PSG - penicillin, streptomycin, and glutamine

RFU – relative fluorescence units

RGB – resolving gel buffer

RSS-rapid size screen

SDS - sodium dodecyl sulphate

SEM - standard error of the mean

SGB – stacking gel buffer

TBE – tris - borate EDTA

TEMED - N,N,N',N'-tetramethylethylenediamine

TX-100 - triton X-10

Vr – reversal potential

Ω - ohm

Contents

1.0 Introduction

Cancer is a global disease that can strike anyone despite the age, race or gender. Current treatment includes the use of chemotherapy, radiation, and surgery, but each is limited in its ability to completely eradicate cancer of the body. Chemotherapy produces an initially good response; however persistent cycles and the non-specificity of drugs used result in chemo-resistance and peripheral toxicity. The application of radiotherapy to treat cancer has been limited by the unpredictable deleterious effect that it has on normal cells. The surgical removal of tumour cells can not completely remove metastatic cells which usually cause a relapse of the disease (Mathew and Verma, 2009). It becomes increasingly clear that in order to successfully treat cancer a specific targeted treatment with permanent effects and minimal damage to healthy cells is required.

The term 'magic bullet' was first coined by Paul Ehrlich to describe toxic compounds that bind to cell receptors. He theorised that toxic compounds or drugs can interact with intended cell structures targeting and attacking pathogens but remaining harmless to healthy tissue. Ehrlich's research on treatment of infectious diseases with drugs derived from the German dye industry indicated that targeting receptors of pathogens absent in the host cells resulted severe effect in patients (Strebhardt and Ullrich, 2008). Current cancer treatments include targeting of tumour associated antigens and carbohydrate antigens expressed by cancer cells by toxins derived from plants and bacteria. Commonly used bacterial toxins include *Pseudomonas* exotoxins, anthrax, and diphtheria toxin. These are incorporated to a selective ligand which results in chimeric immunotoxin proteins to target cancer cells. Immunotoxin proteins are not yet the 'magic bullet' as they have been found to target healthy cells that express target

receptors even at low levels making them nonspecific. They also induce an immune response in human and there is the need for immunosuppressant agents to reduce this response (Mathew and Verma, 2009).

Ohba *et al*. (2009) introduced *Bt* cancer killing parasporins as candidates in the race to finding the 'magic bullet' and the cure for cancer (Ohba *et al*., 2009). Parasporins were first highlighted by Mizuki and colleagues (Mizuki *et al*., 1999; Mizuki *et al*., 2000; Ohba *et al*., 2009; Kitada *et al*., 2005). *Bt* is part of the *Bacillaceae* family (Pigott and Ellar, 2007). They are pathogenic bacteria known to produce toxin molecules that target specific host organisms.

Mizuki *et al*. (1999) claimed that parasporal inclusions of parasporins are different to other *Bt* toxins in the way that these *Bt* toxins are non hemolytic, non-insecticidal and exhibit cytocidal activity on some cancer cells lines (Mizuki *et al.,*1999). There are presently six different parasporin groups, namely PS1, PS2, PS3, PS4, PS5, PS6 (Kitada *et al*., 2005; Yamashita *et al,* 2009; Mizuki *et al*., 1999; Mizuki *et al*., 2000).

 This study uses molecular biology techniques to clone, express, characterise and study the parasporin 3 (PS3) or Cry41Aa/Cry41Ab a *Bt* toxin found in *Bt* crystal inclusions which were reported to have cytocidal activity on some cancer cell lines (Yamashita *et al*., 2005). In doing so this study investigated Yamashita *et al.* (2005) claim that Cry41Aa has

cytocidal activity on cancer cells with the aim to understand and locate regions responsible for Cry41Aa specificity towards certain cancer lines.

1.1 The Bt Bacterium

Bt is found almost everywhere and is capable of surviving in a number of different environments and habitats such as soil, dust of storage containers, and leaves. It has 55 different flagellar serotypes and eight non flagellar biotypes (Schnepf *et al.*, 1998). It is a gram-positive bacterium that belong to the *Bacillus cereus* group which includes *Bacillus. cereus*, *Bacillus. anthracis*, *Bacillus. mycoides*, *Bacillus. pseudomycoides*, and *Bacillus. weihenstephanensi*. (de Maagd *et al.*, 2005; Xu *et al.*, 2014). *Bt* is a spore forming entomopathogenic bacteria that produces toxic parasporal inclusions. It stands out from the other *Bacillus* species by its ability to produce these parasporal inclusions or crystals during sporulation (Xu *et al.*, 2014; Bravo *et al.*, 2011). However, during the vegetative growth phase, some *Bt* strains excrete soluble non-crystal vegetative insecticidal proteins (Vip). Vip *Bt* producing group are relatively small compared to *Bt* that from parasporal inclusion (Bravo *et al.*, 2007).

The crystalline inclusion of *Bt* contains protein protoxins that are either Cry (for crystal) and also known as δ -endotoxin or Cyt (for cytotoxic) proteins (Nagamatsu *et al.*, 2010; de Maagd *et al.*, 2003; Crickmore *et al.*, 1998). The term 'Cry toxin' is defined as a *Bt* derived toxin that can induce a toxic effect on a target organism. Any other protein(s) with noticeable sequence similarity is also termed a Cry toxin. The type and number of

the different protoxins in a given inclusion defines the toxicity profile of a *Bt* strain (Pigott and Ellar, 2007).

The diversity of *Bt* toxins within a strain is possible due to the fact that most of the genes encoding them are found on plasmids as part of structures that include transposable elements. Plasmids hold about 10-20% of the genetic makeup of *Bt* genomes, where many toxic genes are expressed as part of an operon (Baum and Malvar, 2005). *Bt* carries out a number of different molecular mechanisms that allows expression of toxin during its stationary phase of growth (Schnepf *et al.,* 1998, de Maagd *et al.*, 2003). Schnepf *et al.* (1998) describe the toxin as folds of domains that have been recombined to give a novel toxin. The toxins exhibit similar mode of action as mammalian toxins giving weight to the theory that all *Bt* toxins originated from a common evolutionary ancestor.

 The toxins are found within parasporal crystalline inclusions. There are no definitive answers as to why *Bt* would invest energy producing these inclusions, but there are a number of possible reasons. It is thought that in a nutrient limited environment, *Bt* ensures access to future nutrients by undergoing sporulation and producing crystal inclusion that would kill the host insect, providing plenty of nutrients. it is thought that the size and insolubility of parasporal inclusions in neutral pH stops them being leached away from the soil (de Maagd *et al.,* 2003)

There are four Vip families that are toxins secreted into the medium during the vegetative stage of a *Bt* life cycle (Crickmore *et al.,* 2018). These were discovered when *Bt* and *B. cereus* cultures were screened and found to produce protein toxins that were not crystals. Warren *et al.* (1998) describe Vip 1 and 2 as binary proteins (A+B types) that have toxicity towards coleopteran insects (Warren *et al.,* 1998; de Maagd *et al.,* 2003). They both possess the N-terminal signal sequence for secretion and are needed together to be active against some *coleopteran* larvae inducing a binary effect. A possible mode of action is thought to involve Vip2 as the cytotoxic A domain and Vip 1 as the receptor binding B domain. Vip3 is known to target the insect order of some lepidopterans whilst the targets of Vip 4 are still elusive (de Maagd *et al.,* 2003). Vips are thought to cause cell death through different modes of action. Vip 1 is moved to the cell cytoplasm via Vip 2, there it enables ADP-ribosylation of actin. Vip 3 has been observed to cause pores in the planar lipid bilayers of target cells (de Maagd *et al.,* 2003).

1.2 *Bt* toxins

The group encompasses the 3-domain Cry toxins, the binary toxins of *Bacillus. sphaericus*, and the ETX/MTX toxins. These share a certain degree of homology suggesting that they have a common evolutionary origin (de Maagd *et al.,* 2003). The largest category in the group is the 3-domain toxins (Crickmore *et al.,* 2018). Amino acid sequence homology has identified more than 300 *Bt* toxins holotypes which have been organised into 73 *cry* and 3 *cyt* families (Xu *et al.,* 2014). *cyt* toxins share little sequence homology with *cry* toxins. The term 'Cyt toxin' is defined as *'Bt* derived toxin(s) that

exhibits hemolytic activity and other any protein(s) with obvious sequence similarity are termed as such (Schnepf *et al.,* 1998).

A nomenclature that classified Cry crystals based on their amino acid sequence homology and *Bt* host range was first proposed by Höfte and Whitely in 1989 (Höfte and Whiteley, 1989). It was later revised by Crickmore *et al.* (1998)*.* The *Bacillus thuringiensis* Toxin Nomenclature Committee classifies *Bt* toxins on a four-rank system. Toxins are given a primary rank such as *cry1, cry2*, then a secondary rank such as *cry1A, cry2A,* then a tertiary rank such as *cry1Aa, cry1Ab* and quaternary rank such as *cry1Aa1* or *cry1Aa2.* The classification indicates that identified *Bt* toxins sequences respectively share less than 45%, 78%, 95% and ≤100% pairwise identity (Crickmore *et al.,* 1998; Crickmore *et al.,* 2018).

Cry toxins have shown toxicity to specific insect orders that include *Lepidoptera* (butterflies and moths), *Diptera* (flies and mosquitoes) and *Coleoptera* (beetles and weevils); with recent additions of insect orders that include *Hymenoptera, Orthoptera*, *Hemiptera, Isoptera*, *Mallophaga, Thisanoptera*, as well as nematodes and mites (Xu *et al.,* 2014; Nagamatsu *et al.,* 2010; Pigott and Ellar, 2007). Cyt proteins are cytotoxic to vertebrate and invertebrate cells, as well as being haemolytic *in vitro* whilst exhibiting some insecticidal activity in *vivo* (Mizuki *et al.,* 1999). Figure 1 lists *Bt* toxins according to their primary rank (Adang *et al.,* 2014).

Figure 1 List of *Bt* Cry toxins based on their primary rank.
The largest Cry toxins is the 3-domain family shown in blue. The Bin family shown in pink and the ETX/MTX2 family are shown in orange. Other colours indicate unrelated Cry toxins that have not been classified into current family groups. Parasporins are in red font. Taken from Adang *et al.* (2014).

The Cyt toxins are made by mosquitiocidal *Bacillus. sphaericus* strain and organised into four families (Crickmore *et al.,* 2018; de Maagd *et al.,* 2003). *In vitro* studies have shown Cyt toxins to be cytotoxic to a number of different vertebrate cell lines that include erythrocytes (Bravo *et al.,* 2007; Soberón *et al.,* 2013). *In vivo* studies showed that Cyt toxins are mostly toxic to dipteran insects whilst some exhibit toxicity to coleopteran species (Soberón *et al.,* 2013). Cyt toxins have a single domain made up of three-layer alpha-beta proteins with a distinctive fold. Studies have indicated that Cyt toxins play a role in the toxicity of *Bt* Cry toxins by exerting a synergistic effect to overcome Cry mosquitocidal resistance (Bravo *et al.,* 2007; Crickmore *et al.,* 1995; Soberón *et al.,* 2013) Cyt toxins are activated in the susceptible insect gut and result in a 25 KDa protease resistant core. Contrary to Cry toxins which are thought to interact or bind to midgut epithelium cell, Cyt are thought to act on the saturated membrane lipids. Two models have been proposed to explain their mode of action. A pore formation model where Cyt toxins bind to cell membranes and initiate the formation of cation-selective channels in membrane vesicles that subsequently results in colloid-osmotic lysis of the cell. In the detergent effect model, Cyt toxins are thought to aggregate non-specifically on the surface of the lipid bilayer that results in the disarrangement of cell membrane and eventual cell death. Structure of Cyt toxins has shown that they exhibit similar topology with a single domain made up of two outer layers of helix hairpins that coil around α-sheets (Soberón *et al.,* 2013).

With the exception of parasporins, Cry proteins are not known to be harmful to vertebrate or plants cells (Nagamatsu *et al.,* 2010). For many decades, commercialised

pesticide *Bt* spore sprays have been used safely to control agricultural pests (de Maagd *et al.,* 2003). Other species such as *israelensis* exhibit mosquitocidal activity and are used to control disease vectors such as mosquitoes and blackflies. In addition, research on Cry expressing transgenic plants has led to their commercialised use (de Maagd *et al.,* 2003).

The other remaining subgroups within the *Bt* toxin group include the binary and MTX group which are mentioned briefly. The binary (Bin) toxins are found in crystals that contain two separate proteins known as Bin A and Bin B which are homologous to each other and both are required for the recombinant production of toxins in *E. coli*. They are similar to Cry toxins in that they require solubilisation and activation by protease in the insect midgut. Binary toxins are toxic to western corn rootworm. Cry34 and Cry35 are both required in order to cause a toxic effect on target cells and it's requirement for two components, that has referred these toxins as binary insecticidal crystal proteins (de Maagd *et al.,* 2003). Their molecular structure and mode of action is still unclear, but they are thought to cause pores in the target insect gut (Adang *et al.,* 2014).

The MTX Cry family are found in *Bacillus. sphaericus* strains and are toxic to some mosquito species, those made by *Bt,* demonstrate toxicity to some *lepidopterna* and *coleoptern* species (de Maagd *et al.,* 2003) . There are 11 members of the MTX Cry family and they share homology to the cyt toxin of *Pseudomonas aeruginosa*, the epsilon toxin of *Clostridium perfringens*, alpha-toxin of *Clostridium septicum* and aerolysin of *Aeromonas hydrophila* (de Maagd *et al.,* 2003).

The Mtx toxin group are also pore forming toxins and includes two toxins: Mtx2 and Mtx3. Both are related to each other and are distantly related to other *Bt* crystal proteins. Of the two, Mtx 2 demonstrates more amino acid diversity allowing it to change its toxicity between different mosquitos' species (de Maagd *et al*, 2003*)* Figure 2 summaries the protein toxins produced by *Bt* and illustrates their classification according to sequence homology between them. Figure 3 shows the insect orders susceptible to *Bt* δ endotoxins (Palma *et al.,* 2014).
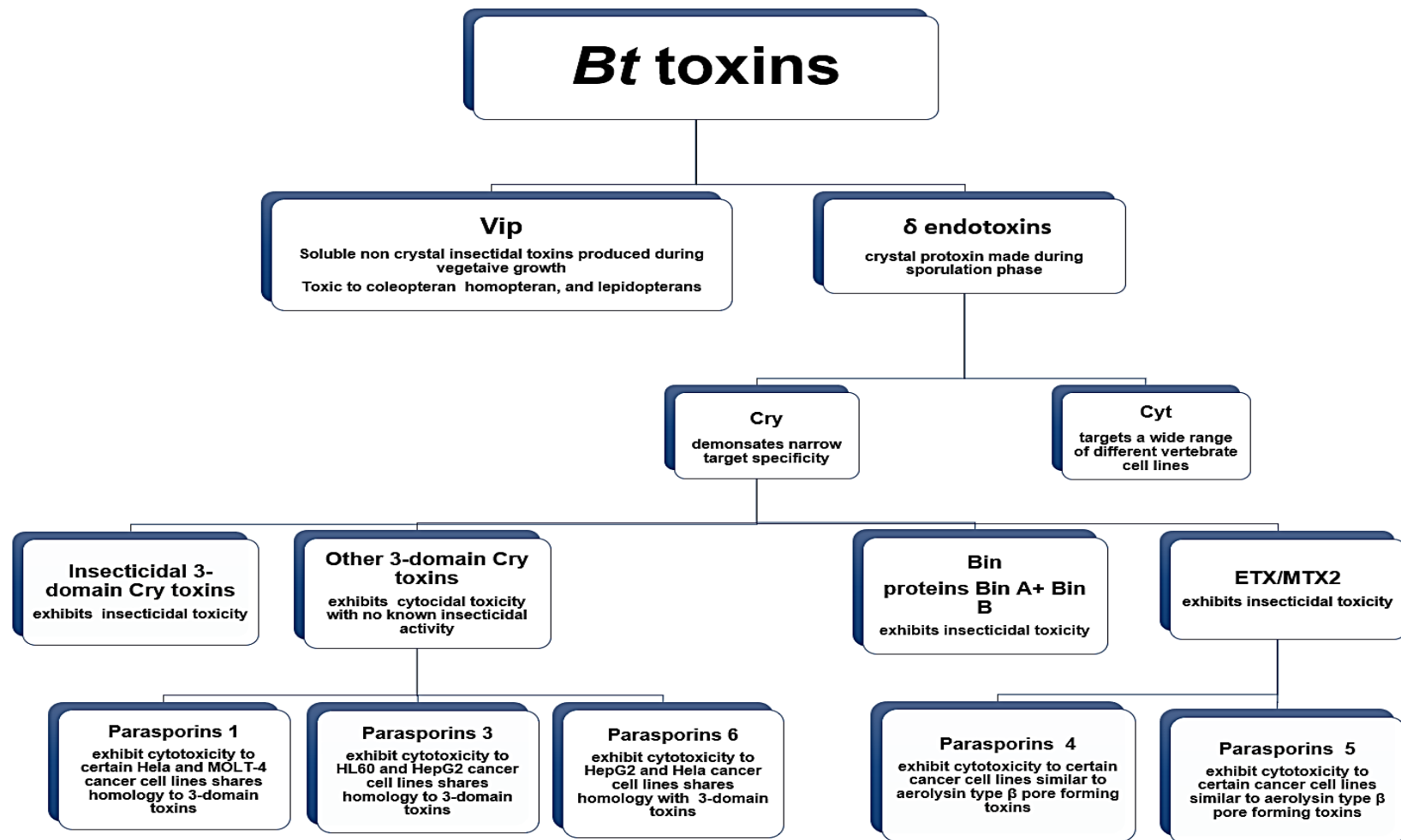
Figure 2 Classification of parasporin toxins produced by *Bt*
Figure summarises the parasporin toxins produced by *Bt.* These are classified according to sequence homology to other protein toxins produced by *Bt*. The Cry proteins are comprised of phylogenetically different families: insecticidal 3-domain, other 3-domain Cry toxins, Bin, and ETX/MTX2.
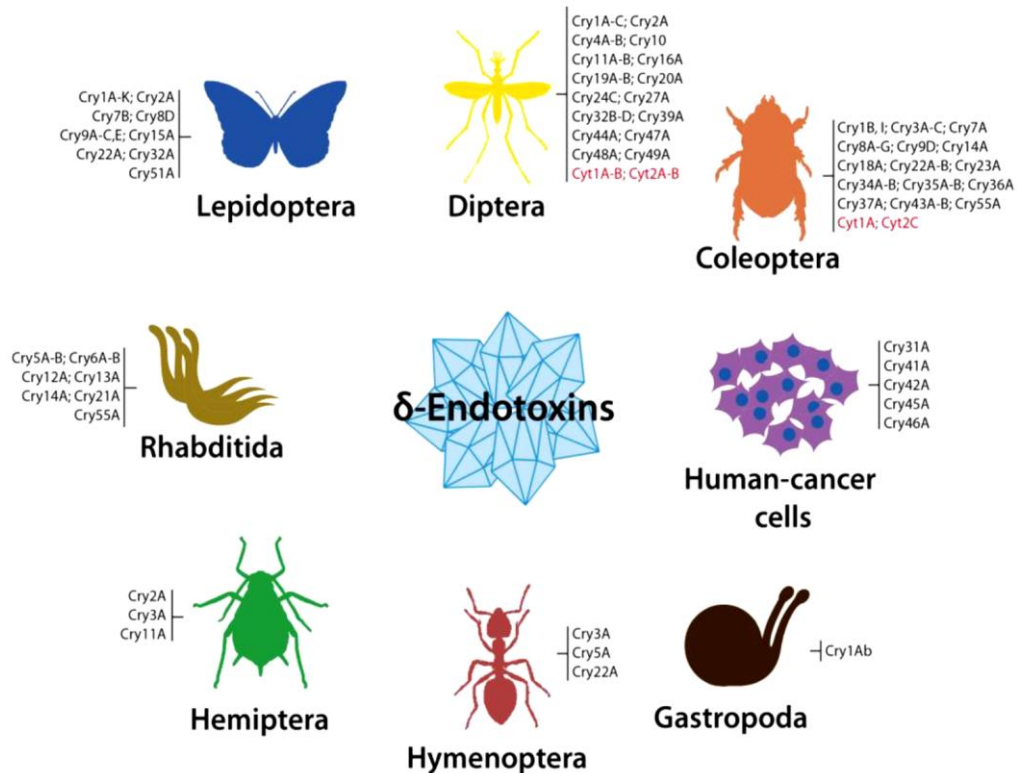
Figure 3 Diversity of hosts and *Bt* toxins organised based on target specificity
Figure illustrates both Cry and Cyt *Bt* δ endotoxins and type of insect orders that they kill. The toxins are grouped according to target specificity. Taken from Palma *et al.* 2014.

1.2.1 3-domain Cry toxins

Cry toxins belonging to the three-domain Cry toxin family, demonstrate differences in their amino acid sequences but they all share a conserved three-domain structure. The 3-domain toxins are described as pore forming toxins that can kill insect orders Lepidoptera, Diptera, Coleoptera, Hymenopters, and some nematodes. The protoxins have a usual length of ˜ 130 KDa or the smaller size of ˜ 70 KDa (de Maagd *et al.,* 2003). These smaller 3-domain protoxins do not have an extended C-terminal region(Schnepf *et al.,* 1998).

There are four theories that hypothesise Cry mode of action, these will be discussed in section 2.4. It is generally accepted that 3-domain toxins are proteolytically activated by

proteases in the midgut of insects; the crystalline protoxin is solubilised in alkaline conditions allowing cleavage of the N- and C-terminal by gut proteases that results in an active protease resistant core of ˜60 KDa. The active toxin binds to the brush-border membrane of the insect gut epithelium. It can do so by the affinity it has for receptors present on the membrane of target cells, which also defines the specificity that Cry toxins exhibit to insect orders. Once bound, part of the toxin forms ion channels and aggregates in such a way that causes formation of larger pores within the membrane. Osmotic imbalance causes the cells to swell and subsequently lysis. Within min the insect stops feeding and eventually dies (Baum and Malvar, 2005, Nagamatsu *et al.,* 2010; Katayama *et al.,* 2005; Mizuki *et al.,* 1999; Ohba *et al.,* 2009; Schnepf *et al.,* 1998).

The 3-domain Cry toxins are composed of three distinctive domains, domain I, II, and III. Each of the three domains is thought to play a role in the mode of action of Cry toxins. Multiple-sequence alignments of different Cry toxins have highlighted the existence of five conserved blocks located in the active toxic core of the protoxins which encompasses the three domains. These blocks are typically found in the active toxin post activation by midgut proteases (Höfte and Whiteley, 1989) and are thought to play an important role in toxin stability and function and are located in the N-terminal of the larger protoxins ˜130 KDa (Pigott and Ellar, 2007). In addition to the five conserved blocks, there are three more conserved block sequences away from the active core and towards the C-terminal end of the protoxin (Schnepf *et al.,* 1998). The C-terminal of the protoxin is thought to be involved in crystal formation and contains conserved blocks 6,7, and 8 as shown in figure 5 page 29.

Domain I contains conserved block 1 and is thought to take part in pore formation and exhibits similarities to other pore forming toxins. Li *et al*. (1991) first described domain I of Cry3Aa as an α-bundle, where amphipathic helices (bigger than 30 Å) surround a central helix (Li *et al.,* 1991). The helices are well charged with polar, and hydrophobic residues that seem to aim towards the central helix. The polar residues are present in interhelical spaces as hydrogen bonds or as part of salt bridges. These characteristics and the resemblance to the pore forming domain of colicin (an antibiotic released by bacteria to kill other bacteria of the same species) has proposed domain I as a key factor of pore formation in Cry toxins (Pigott and Ellar, 2007).

Domain II is thought to interact with insect gut receptors and demonstrates similar properties to carbohydrate binding proteins. Three antiparallel β sheets are bundled to form a β-prism. The conserved block 2 is partly found in domain I and in the first β sheet of domain II. Two of the β sheets are made of four strands in a Greek key motif. The third is located against domain I and arranged in a similar Greek key motif with shorter stands.

Of the three domains, domain II exhibits the biggest structural diversity particularly the apex loops which differ in length, structure, and sequence. Exposed loops 1,2,3, and α8 have been identified as regions that affect the specificity of Cry toxins and modifications here have been known to alter the toxin's ability to interact and kill target insects (Bravo *et al.,* 2013). In addition to this, the length of β strands also vary from one Cry toxins to another. It is this diversity that is thought to determine toxin specificity. There appears

to be a structural similarity between domain II and other β prism proteins such as plant lectins, vitelline, *Maclura pomifera* agglutinin etc, which are known to have binding properties. Likenesses to complementary determining region of immunoglobulins, studies have also supported the theory that domain II loops are involved in receptor recognition (Pigott and Ellar, 2007).

Domain III is thought to be involved in receptor binding and pore formation and is similar to carbohydrate binding protein domains. Ithas a β-sandwich structure that contains conserved blocks 3, 4, and 5 in one of its three buried stands (de Maagd *et al.,* 2003; Schnepf *et al.,* 1998; Nagamatsu *et al.,* 2010). Two long loops appear to reach out and interact with domain I. Domain III demonstrates less diversity compared to the other two domains (Pigott and Ellar, 2007). It is thought that domains II and III are involved in receptor binding and specificity (Lee *et al.,* 1996; Pigott and Ellar, 2007).

X-ray crystallography has helped to reveal the crystal structure of the following Cry toxins Cry1Aa, Cry1Ac, Cry2Aa, Cry3Aa, Cry3Bb, Cry4Aa, Cry4Ba, Cry5Ba, and Cry8Ea (Li *et al.,* 1991; Grochulski *et al.,* 1995; Morse *et al.,* 2001; Galitsky *et al.,* 2001; Boonserm *et al.,* 2006; Guo *et al.,* 2009; Hui *et al.,* 2012; Boonserm *et al.,* 2005). The toxins share the 3-domain structure but differ in amino acid sequence and insect specificity. Figure 4 illustrates examples of the 3-domain structure of Cry toxins (Pigott and Ellar, 2007). Figure 5 illustrates the five sequence blocks typical of 3-domain toxins which were first described by Höfte and Whiteley (1989), as well as the remaining three sequence blocks later added by Schnepf *et al.* (1998). Upon treatment with protease, conserved blocks

6-8 are usually lost leaving a protease resistant core with the five conserved blocks

(Schnepf *et al.,* 1998; Höfte and Whiteley, 1989; Palma *et al.,* 2014).
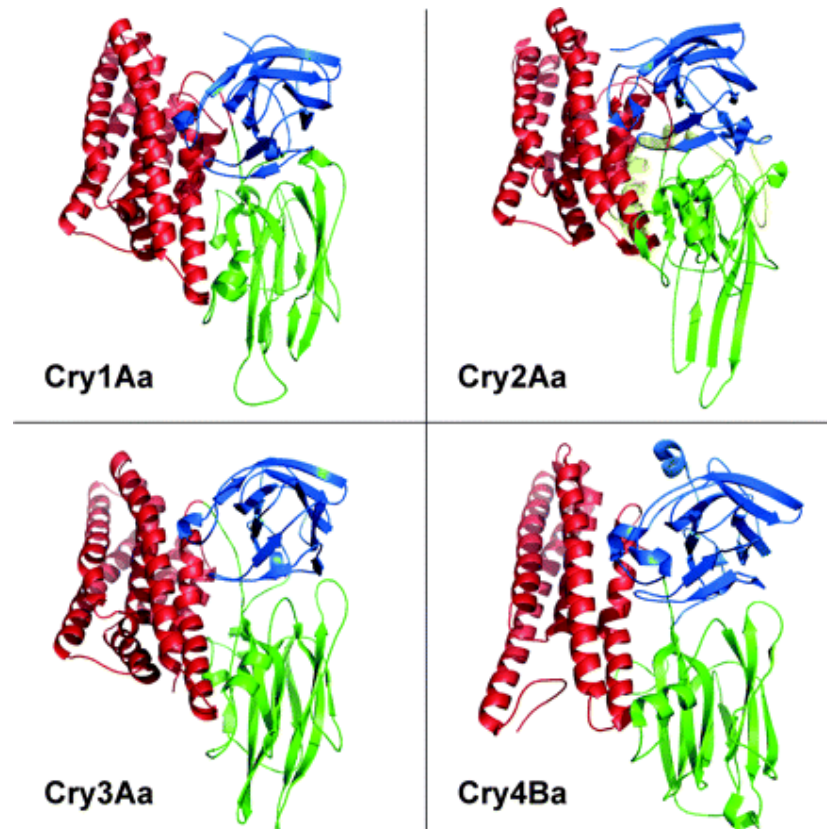
Figure 4 Crystal structure of 3-domain Cry toxins.
Figure shows the crystal structure of 3-domian toxins: Cry1Aa, Cry2Aa, Cry3Aa and Cry4Ba. Domain I is shown in red, domain II is shown in green, and domain III is shown in blue. Taken from Pigott and Ellar *et al.* 2007.
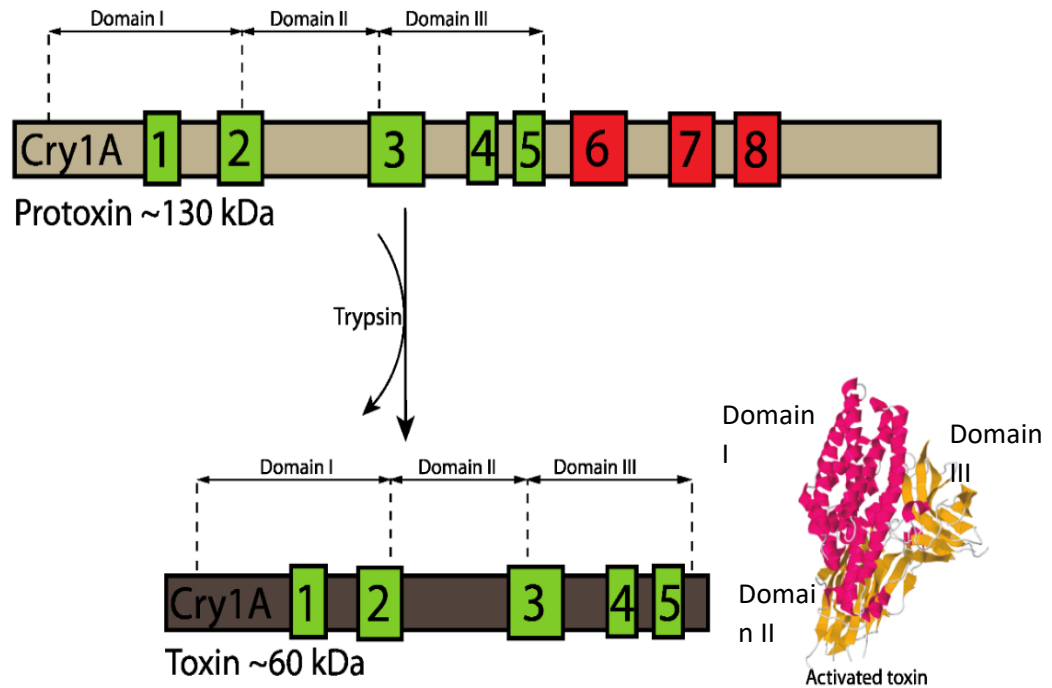
Figure 5 conserved sequence block of 3-domain toxins.
Figure illustrates the 5 conserved sequence blocks of 3-domain toxins of Cry1A in green and the later 3 sequence blocks in red. Upon treatment with trypsin enzyme blocks 6 to 8 are degraded leaving a protease resistant core of ˜60 KDa. Adapted and taken from Palma *et al.* (2014)

The number of different insecticidal proteins produced by *Bt* indicates that their genes are affected by selective evolutionary pressures that lead to the development of a wide range of *Bt* targets (Pigott and Ellar, 2007). Bioinformatic investigations have revealed that each one of the 3-domains may be evolved at a different rate and therefore affect target specificity (Bravo, 1997).

1.2.2 Parasporins

The number of *Bt* strains with no known insecticidal activity far outnumbers insecticidal *Bt* strains (Yasutake *et al.,* 2005; Uemori *et al.,* 2005; Mizuki *et al.,* 1999; Ohba *et al.,* 2009; Schnepf *et al.,* 1998; Nagamatsu *et al.,* 2010; Kitada *et al.,* 2005). This finding prompted a review of *Bt* toxin action against a range of invertebrate and vertebrate cell lines. For the first time, Mizuki *et al.* (1999) reported a non-insecticidal parasporal inclusion, isolated from a *Bt* strain, that is cytocidal to human cancer cells (Mizuki *et al.,* 1999; Ohba *et al.,* 2009). Research has since discovered a number of different *Bt* proteins with a range of cytotoxicity against mammalian cell lines (Ohba *et al.,* 2009).

This has led to a new category of *Bt* toxins known as the parasporins which are genealogically heterogeneous to Cry proteins. The Committee of Parasporins Classification and Nomenclature was formed and classified parasporins into six families: PS1, PS2, PS3, PS4, PS5, and PS6 (Ohba *et al.,* 2009; Chikawa *et al.,* 2008; Okumura *et al.,* 2010). Similar to the ranking of Cry toxins, parasporins were ranked according to their primary amino acid sequence and given a four rank name according to the level of sequence identity to known toxins as summarised in table 2 (Crickmore *et al.,* 1998).

They were purified from independent *Bt* strains and isolated from soil in Japan taken from Hiroshima, Fukuoka, and Tokyo (Kitada *et al.,* 2005) as well India (Poornima *et al.,* 2010), the Caribbean islands (Gonzalez *et al.,* 2011), and Malaysia (Nadarajah *et al.,* 2006). Parasporins differ from Cry and Cyt proteins, as they do not exhibit any insecticidal activity. Despite being non-haemolytic, they are cytotoxic to some

mammalian cell lines (Yasutake *et al.,* 2005; Katayama *et al.,* 2005; Ohba *et al.,* 2009; Kitada *et al.,* 2005).

Parasporins share low but significant sequence similarity with other Cry toxins (Ohba *et al.,* 2009). They resemble Cry proteins in that they all require proteolytic processing. The activated form of Cry proteins have about 600 amino acids and some have the 3-domain structure (Nagamatsu *et al.,* 2010).

Research has indicated that parasporins differ in their cytotoxicity mechanism and type of target cells. Katayama *et al.* (2005) argue that parasporins differ in their mode of action to Cry proteins. They propose that PS1 is not a membrane binding protein. Membrane binding protein usually cause irreversible permeability of ions in the membrane that lead to depolarization and change in membrane potential which subsequently results in cell death. Instead, they propose that PS1 induces Ca^{2+} influx in cells without a change in membrane potential. This increase disrupts Ca^{2+} homeostasis or other cell pathways that subsequently lead to apoptosis as a result of a drop in cellular protein and DNA synthesis (Yamashita *et al.,* 2005; Kitada *et al.,* 2005). In other words, it is not a pore forming toxin.

Mizuki *et al.* (2000) argue that cytotoxicity mechanisms of Parasporins are not fully understood and it is uncertain if cell death is caused by induced apoptosis alone. Uemori *et al.* (2009) confirmed the ability of PS1 to discriminate between healthy uterine smooth cells and uterus cervix cancer cells. Furthermore, the general characteristic of

parasporins of having specificity for cancer cells lead to proposing them as a potential 'magic bullet' in treating cancer (Ohba *et al.,* 2009).

Concerns regarding *Bt* specificity are a current debate. Uncertainties regarding *Bt* toxin synergism, efficiency and selectivity are being observed in their application as pest controls agents (Then, 2010). Mizuki *et al.* (2000) argue that *Bt* toxins are also cytocidal to normal mammalian cells line, highlighting the need for refinement of their specificity in order to reduce the potential danger. Yamashita *et al.* (2005) also observed 'low toxicity' to four non-cancerous human cell line by Cry41A proteins. The cytocidal activity of parasporin toxins varies in range and intensity, as well as the range of target cells. For example, PS4 is preferentially toxic to leukaemia T cells, however it shows low toxicity to normal T cells, whilst exhibiting no toxicity at all to uterus cervix cancer cells (Saitoh *et al.,* 2006)

1.2.3. Parasporin 1 /Cry31Aa1

Parasporin 1 or Cry31Aa1 shares less than 25% sequence identity with *Bt* toxins, however, it does contain the five conserved blocks typical of 3-domain Cry toxins (Mizuki *et al.,* 2000). The crystal structure was resolved at 1.76 Å resolution and was revealed to have a 3-domain structure typical of most insecticidal Cry toxins. An N-terminal region similar to Cry2Aa was also detected (Katayama *et al.,* 2005; Kitada *et al.,* 2005; Akiba and Okumura, 2016).

Upon proteolytsis two active polypeptides of 15 and 56kDs were formed and were toxic towards HeLa, HL60, MOLT-4 and HepG2 cell lines (Katayama *et al.,* 2007). An alternative mode of action has been proposed for Cry31Aa1 where interactions with Hela cells is thought to induce an apoptotic signalling pathway that causes a rapid influx of $Ca2^+$. Cell viability decreased progressively and morphological indicators of cytopathology such as cell swelling were observed 8-10 h post incubation with Cry31Aa1 despite a dose of 10 µg/mL. The cell membrane remains intact and no change in cell permeability was observed (Katayama *et al.,* 2007). This is quite unusual for parasporins as membrane permeability is often observed with other parasporins (Ohba *et al.,* 2009; Yamashita *et al.,* 2005).

In Cry31Aa1 susceptible cells the membranes were impermeable to DNA binding dye molecules and the cells remained polarized. Furthermore, no cytosolic markers such as LDH were observed even after 4 h of incubation with Cry31Aa1. However, within min, intracellular calcium probe fura-2 detected a sudden rise in intercellular $Ca2^+$. It is thought that $Ca2^+$ was depleted from the extracellular space as investigation with low $Ca2^+$ levels correlated with reduced toxicity (Katayama *et al.,* 2007). A number of $Ca2^+$ influx inhibitors were applied to investigate if they affect $Ca2^+$ levels and thus toxicity of Cry31Aa1. Only suramin which inhibits G protein coupled receptors (GPCR) affected $Ca2^+$ levels. Elevated levels of caspase 3 and PARP cleavage in treated cells were indicative of an activated apoptotic signalling where DNA and protein synthesis were prevented.

The findings led to the proposal that Cry31Aa1 causes cytotoxicity by the activation of GPCR, calcium influx and apoptosis due to ineffective calcium homeostasis without any form of pore formation (Katayama *et al.,* 2007).

Contrary to this, is a study by Narvaez *et al.* (2014) where they argue that parasporin 1Aa2 or Cry31Aa2 has 94% sequence identity with Cry31Aa1 and has been observed to form pores in artificial membranes. In addition, it was found to induce calcium oscillations despite the lack of extracellular calcium and normal cell line HEK 293 (Gabriel Narvaez *et al.,* 2014). It was suggested that Cry31Aa1 and Cry31Aa2 have the potential to form pores as evident with the formation of ion channels in artificial membranes.

In response to this, further studies investigated the occurrence of an additional N-terminal region which locks domain I and stops it from forming pores (Akiba and Okumura, 2016; Akiba *et al.,* 2005). Katayama *et al*. (2011) carried out an investigation in search of the receptor for Cry31Aa1 and discovered beclin-1 a tumour suppressor protein. Anti-beclin-1 antibody repressed toxin binding and toxicity. It was patented and reported as a receptor for Cry31Aa in Hela cell lines (Katayama *et al.,* 2011).

1.2.4. Parasporin 2 /Cry46Aa

Parasporin 2 or Cry46Aa1 is a very toxic protein and shares very little amino acid identify with Cry toxins. The crystal structure of the activated protein was resolved and revealed an elongated form of 115 x 30 x 29 Å. Structurally it is similar to the ETX/MTX2 Cry family

and protein in the Toxin_10 family as well as hydralysin (Akiba *et al.,* 2009; Xu *et al.,* 2014).

It was mainly tested on HepG2, but also has shown toxicity to a number of different cancer cell lines such as MOLT-4, and Jurkat (Kitada *et al.,* 2006). In addition to being tested on cancer cell lines, it was also investigated for its effects on cancer tumour slices from differentiated hepatocellular carcinoma and colon cancer. It was found to cause cell swelling in non -neoplastic cancerous cells (Kitada *et al.,* 2006; Ito Akio *et al.,* 2004).

Kitada *et al*. (2006) demonstrated the toxin interaction through a number of experiments. Immunofluorescence visualised the plasma membrane of HepG2 cells where Cry46Aa1 caused cell swelling, blebbing and the breakdown of organelles and cytoskeleton. The release of LDH and PI staining cell demonstrated the instant damage to the cell membrane. This was confirmed by osmoprorective PEGs analysis which measured Cry46Aa1 induced pores with a diameter of 3nm (Kitada *et al.,* 2006).

In addition to pore formation, Cry46Aa1 was found to induce caspase activation and DNA degeneration in HepG2 cells. Further investigation by Ito Akio *et al.* (2009) and others indicated that Cry46Aa maintained cytotoxicity even after the application of a caspase inhibitor and suggested that its main mode of action is by pore formation (Ito Akio *et al.,* 2004; Akiba *et al.,* 2009).

Interactions with lipid rafts in HepG2 cells required GPI-anchored proteins for oligomerization and toxicity (Akiba *et al.,* 2009; Okumura *et al.,* 2011). Its structural resemblance to proteins that possess additional β trefoil motif in the head domain and aerolysin type β pore forming toxins gives weight to Cry46Aa1 being a pore forming toxin (Akiba *et al.,* 2009; Sher *et al.,* 2005).

1.2.3 Parasporin 3 / Cry41Aa, Cry41Ab

Parasporal inclusions were isolated from *Bt strain* A1462, parasporal inclusions were successfully solubilised in alkaline conditions, and two 88 KDa proteins were recovered. The proteins were respectively named Cry41Aa (PS3Aa) and Cry41Ab (PS3Ab).

They were treated with proteinase K and analysed on SDS PAGE gel where each revealed a 64 KDa major protein and a minor 80 KDa protein. The proteolytically active proteins were cytocidal activity against HepG2 (hepatocyte cancer) and HL60 (myeloid leukaemia cancer) cells. The two proteins were respectively named Cry41Aa and Cry41Ab.

The proteins were expressed from three genes ORF1, OFR2, OFR3. ORF1Aa and OFR1Ab proteins are similar in weight and show 87% identity. ORF2Aa and ORF2Ab are 88% homologous and of analogous weight. OFR3Aa and ORF3Ab have 99% sequence homology and are of similar weight. There were no transcriptional terminator sequences between ORF1 and ORF2 or ORF2 and OFR3.

OFR2 contained the five conserved block regions common to three domain Cry toxins, and ORF3 contained conserved Cry block regions 6, 7, and 8 also found in some insecticidal Cry toxins. The N-terminal amino acid sequences of the proteolytically resistant purified 64 KDa proteins were found in ORF2Aa and ORF2Ab. The C terminal sequence of both ORF2's resembles the C*lostridium botulinum* hemagglutinin HA-33 sequence.

Yamashita *et al*. (2005) attempted to express the proteins using only ORF2 genes but observed that *Bt* could not form parasporal inclusion. This suggests that the expression of ORF3 may have a role in protein conformation and/ or packaging. It is important to note that alkali only treated proteins were cytocidal and proteolytic activation of both Cry41Aa and Cry41Ab was required for their characteristic toxicity towards HepG2 and HL60 cancer cell line.

The study further suggests that Cry41Aa and Cry41Ab may have a specific receptor on susceptible cancer cells. The findings from cytotoxicity assay and microscope observations suggest that both Cry41Aa and Cry41ab kill slowly and are both likely to have a 3d-Cry structure common to some insecticidal Cry toxins (Yamashita *et al.,* 2005).

Yamashita *et al*. (2005) screened 1744 Bt parasporal protein against 124 human cell lines, of those Cry41Aa was toxic to HL60 and HepG2 cell lines. The EC50 for HepG2 cell was 1.86 µg/mL (2.4 µM/ml) and 1.25 µg/mL (1.6 µM/ml) for HL60 cell lines. Cell viability and membrane damage assay have suggested that Cry41Aa is more toxic than Cry41Ab

as indicated by the dramatic decrease in cell viability and an increase of LDH efflux. In addition to MTT viability cell assay trypan blue staining highlighted morphological changes of the cells during their 24-h incubation with Cry41Aa and Ab toxins. Cells were observed to swell within the first hour followed by membrane damage in the hours that followed, as shown in figure 6 (Yamashita *et al.,* 2005).

list of mammalian cell lines tested against proteinase K activated Cry41Aa and Cry41Ab.

Cell		Origin	EC$_{50}$(µg/mL)	
			Cry41Ab	Cry41Aa
Human				
	MOLT-4	Leukemic T cell	>10	>10
	Jurkat	Leukemic T cell	>10	>10
	HL60	Myeloid leukaemia	1.32	1.25
	HeLa	Uterus cervix cancer	>10	>10
	TCS	Uterus cervix cancer	>10	>10
	Sawano	Uterus cancer	>10	>10
	HepG2	Hepatocyte cancer	2.80	1.86
	A549	Lung cancer	>10	>10
	CACO-2	Colon cancer	>10	>10
	T cell	Normal T cell	>10	>10
	UtSMC	Normal uterus	>10	>10
	HC	Normal hepatocyte	>10	>10
	MRC-5	Normal lung	>10	>10
Simian				
	Vero	African green monkey kidney	>10	>10
	COS-7	African green monkey kidney	>10	>10
Murine				
	NIH3T3-3	Mouse embryo	>10	>10

Table 1 The list of mammalian cell lines tested against proteinase K activated Cry41Aa and Cry41Ab.
MTT cell viability assay was applied post a 24-h incubation with toxins. Data for HL60 and HepG2 cell lines highlighted in red. Taken and adapted from Yamashita *et al.*2005.
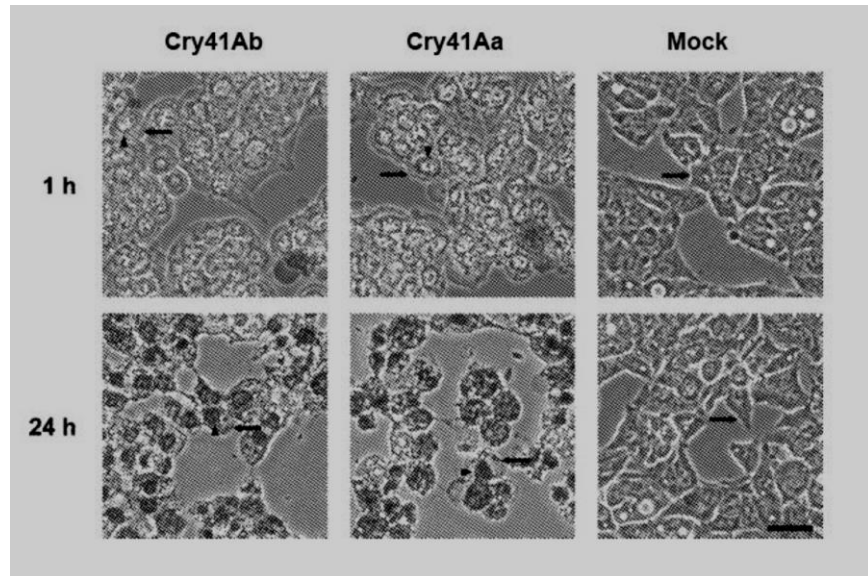
Figure 6 Morphological changes of HepG2 cells post incubation with Cry41Aa.
Microscope images were taken of Trypan Blue stained HepG2 cell line after incubation with Cry41Aa and Cry41Ab at 1 h and 24-h post-treatment. Adapted from Yamashita *et al.* (2005).

The Cry41Aa gene was expressed by Krishnan, 2013 (Krishnan, 2013). Primers were designed to amplify the split toxin as ORF2 and ORF3. ORF2 encodes the active toxin with the five conserved Cry blocks typical of 3-domain Cry toxins, as well the N terminal and C-terminal region located within the ricin domain of the toxin as shown in figure 7.

ORF3 encodes the remaining conserved Cry block 6,7, and 8. These are usually lost after protease treatment but are thought to play a role in crystal formation and the packaging of protoxin in *Bt* cell (Krishnan *et al.,* 2017). In addition to the ORFs restriction sites for BamHI, Xhol, and Xbal were also subcloned into Bt shuttle vector pSVP27A. expression is driven by a strong *Bt* cyt1Aa promoter (Crickmore and Ellar, 1992).

The vector was introduced into *Bt*4D7 to express cry41Aa. Upon treatment with proteinase K Cry41Aa is cleaved at position 60 after an alanine which is necessary for the toxicity of Cry41Aa (Yamashita *et al.,*2005) However, N-terminal sequence analysis has indicated that trypsin cleaved Cry41Aa at position 63 after an arginine (Souissi, 2018, Krishnan *et al.,* 2017).

```
MNQNCNNNGYEVLNSGKGYCQPRYPFAQAPGSELQNMGYKEWMNMCTSGDPTVLGEGYSA
DVRDAVITSINIASYLLSVPFPPAGVAAGILGALLGLLWPTNTQAVWEAFMNTVEALINQ
KLDEYARSKAISELNGLKNVLELYQDAADDWNENPGDLRNKNRVLTEFRNVNGHFENSMP
SFAVRNFEVNLLPVYAEAANLHLLLLRDAVKFGEGWGMSTDPGAERDDMYRRLRSRTEIY
TDHCVNTYNQGLQQAKSLQANVSDYSRYPWTQYNQSGGFSYREAKGEYRGTENWNLYNAF
RRDMTILVLDIIAQFPTYDPGLYSRPVKSELTREVYTDIRGTTWRSDANLNTIDAIENRM
VGSRQLQLFTWLTEMKFYIRNTGSITSYTHGDLMVGLEKKIRKTNDNDQWLPLEGQNTSY
TRIDRPGIELGKNYWYYARTQQWFETRLLQLWANTDVLSLNAGTVGNEFWVRDVPDYRNI
YARSTRNHFIENHRLSWIKFEPVRDNCPFAWPGYKQLSALLFGWTHNSVDLNNIISQYRI
TQIPAVKAYWNRGAFSVIRGPGSTGGNLVQLGTGGEVSVKVRPEQTGSDWYRVRIRYAAG
SRGRLNVKKYVSSIHASVTYDYNMTMSSSTQGTYNSFQYLDVYNFRLAEPEFEVWLTNES
GGPIWIDKIEFIPLSPIPELPVYPGTYQIVTALNNSSVVTSEEFCMGIGLTTRCGVNLWS
NNGNTLQKWRFVYNGDQNAFQIKSTPNEDLVLSGSNSGTSVTAETNQNRPNQYWLIEEAG
NGYVYLRSKGNPNLVLDVAGTSTANGTNIILWNYNGSTNQKFKLS
```

Encoded by ORF2

```
MNYNVTKAREAVQALFSNPTTLQLKVTDHHVNQVARLVECIADQIHPKEKMCLLDQVKLA
KRLSRERNLLNYGDFESSDWVGTDGWNVSTNVYTVADNPIFKDHYLNMPSANNPILSDKI
FPTYAYQKVEESRLKPYTRYIVRGFVGSSKDLEILVARYDKEVHKRMNVPNDIIPTSPCT
GEPVSQPTPYPVMPSNTMPQDMWCNPCGNGYQTAAGMMVQSTGMMCQDPHEFKFHIDIGE
LDMERNLGIWIGFKVGTTEGMATLDNIEVVEVGPLTGDALTRMQKRETKWKQKLTEKRMK
IEKAVQIARDAIQTLFTCPNQSCLQSAITLQNILRAEKLVQKIPYVYNQFLQGVLSAVPG
EAYAYDIFQQLSDAVATARALYNQRNVLNNGDFSAGLSNWNGTEGADVQQIGNASVLVIS
DWSASLSQHVYVKPEHSYLLRVTARKEGSGEGYVTISDGTEENTETLKFMVGEETTGATM
STIRSNIRERYNERNMATPDPDAYGGTNGYASNQNMVNYSSENYGMSAHSGNNNMNYQSE
SFGSKPYGDGNSMINGSSNNYEANGYPGNNNINDQSENYGANAYSSNNMNYQSESSGFTP
YGDENNMTNYPSNNYEMNPYSSDMNMSMNRGSDCGCGCSANAYPGGNMMMNNYSSSTYEM
NTYPSSTNMTNHQGMGCGCHYSTNEYPMIEENIPDFSGYVTKTVEIFPETNRVCIEIGET
AGTFMVESIELIRMDCE
```

Encoded by ORF3

Figure 7 Amino acid sequence of Cry41Aa protoxin encoded by ORF2 and ORF3
Figure shows the amino acid sequence for split toxin Cry41Aa as encoded by ORF2 and 3. The conserved Cry block are underlined in red. The N-terminal is underlined in black. The exta loop distinctive to Cry41Aa is underlined in dashed lines. The C-terminal/ricin domain are double underlined. Domain I is highlighted in grey, domain II in yellow and domain III in blue.

1.2.4 Parasporin 4/ Cry45Aa1

Okumura *et al*. (2008) first described Cry45Aa1 as a pore forming parasporin that exhibits toxicity towards CACO-2, Sawano and MOLT-4 cells (Okumura *et al.,* 2008).

 It shares little homology to Cry toxins and does not have the five conserved blocks typical of 3-domain toxins. Structurally it is similar to aerolysin type β pore forming

toxins as a consequence of the high level of β structures as measured by CD spectrum. It has a low identity to other β pore forming toxins such as the α toxin from *C.septicum*, epsilon from *C. perfringens*, and aerolysin from *A. hydrophila* (Okumura *et al.,* 2011).

In MOLT-4 cell assay, cells demonstrated cytopathic characteristics within 10 min of incubation with Cry45Aa1; the nucleus fragmented within the hour, and cells completely erupted after a 24-h period. Membrane damage assay detected the formation of large pores that allowed the passage of 70 KDa dextrans and LDH efflux (Okumura *et al.,* 2011).

Okumura *et al*. (2013) discovered additional proteins produced by the same A1470 *Bt* strain, the protein PS 2Aa2 which are almost identical to PS2 (differ by 4 amino acids) was observed to kill MOLT-4 cells with an EC_{50} of 0.47 μg/mL. Interestingly, *Bt* produced activated Cry45Aa1 was more toxic (EC_{50} = 0.13 μg/mL) than the recombinant activated Cry45Aa1 protein (Okumura *et al.,* 2011; Saitoh *et al.,* 2006).

Further investigations on Cry45Aa1 showed that it was activated in acidic conditions and was cleaved by pepsin. This alternative activation did not just increase toxicity of the protein, but it was reported to yield more proteins than through typical alkali conditions (Okumura *et al.,* 2008; Okumura *et al.,* 2014).

Further investigation by Okumura saw the application of Cry45Aa in mice trials. Here, the ability of Cry45Aa to remain stable in low acidic conditions was exploited. It was assumed that the acidic conditions and the presence of pepsin in the digestive system of mice would solubilise and activate Cry45Aa1 which was administered orally. The study concluded that protoxin Cry45Aa1 was indeed activated in the digestive tract of mice and impaired kidney function (Okumura *et al.,* 2014).

1.2.5 Parasporin 5 /Cry64Aa1

This toxin was first described by Ekino *et al.* (2014) as a parasporin with some sequence identity to *Bt* Cry toxins and aerolysin type β pore forming toxins. It is 34 KDa protoxin and was cleaved by proteinase K at the C-terminal to produce a 30 KDa toxin. It exhibits cytotoxicity to MOLT-4 (EC_{50} = 0.1µg/mL), HepG2, TCS, Hela, Sawano and two monkey kidney cell line COS7 and Vero. MOLT-4 cells were observed to swell within an hour of incubation with the toxin (Ekino *et al.,* 2014).

1.2.6 Parasporin 6/Cry63Aa

Cry63Aa is cytotoxic protein with a protoxin of 84 KDa Which upon proteolytic processing produced two core resistant cores of 14 and 59 KDa. These were found to be active against HeLa and HepG2 cells. Structurally, the protoxin N-terminal shares identity with Parasporin 1, however this is lost during protease treatment. The resulting protein core shares 56% sequence identity with Cry2A toxin and in particular the region with the five conserve blocks of Cry toxins (Nagamatsu *et al.,* 2010).

Summary of Parasporin members and their toxins

	PS1/Cry31Aa1	PS2/Cry 46Aa1	PS3/ Cry41Aa1 Cry41Ab1	PS 4/ Cry45Aa1	PS 5/ Cry64Aa1	PS 6/Cry63Aa1
Protoxin (KDa)	81	37	88	31	33.8	84
Toxin (KDa)	15 and 56 heterodimer	30	64	27	29.8	14 and 59
In vitro activated by	Trypsin	Proteinase K	Proteinase K	Proteinase K	Proteinase K	Trypsin
Proteolytic activity site	Two sites in N-terminal	N and C-terminal	N-terminal and possibility in C terminal	C-terminal	C-terminal	N-terminal
Cell death	Ca2$^+$ influx	Cytolysis	Unknown-potential pore formation	unknown	Unknown-potential β pore forming toxins	Unknown-potential pore formation
Receptors	Belclin 1	GPI-Proteins	unknown	unknown	unknown	unknown
Structural homology to Cry toxins	Crystal structure resolved to reveal homology to 3- domain Cry structure and	Crystal structure resolved to reveal homology to aerolysin type β pore forming toxins	Potentially has a 3-domain structure	Potentially similar to aerolysin type β pore forming toxins	Potentially similar to aerolysin type β pore forming toxins	Potentially has a 3-domain structure
Conserved Cry block	five conserved Cry blocks detected in 56 KDa -low homology in block 3	No homology	five conserved Cry blocks detected in 64 KDa	No homology	No homology	five conserved Cry blocks detected in 59 KDa
Sequence homology to Cry toxins	Low, less than 25%	Low less than 25% and other pore forming toxins 41% sequence identity with hydralysin from *Ch. viridissima*	Low, less than 30% 34% sequence identity with Cry3Ba ricin type domain detected	Low, less than 30% and other pore forming toxins	Homology to ETX/MTX2	56.4% homology to 3-domain toxins
Susceptible cancer lines (EC50µg/mL)	HeLa (0.12), HL-60 (0.32), MOLT-4 (2.2), HepG2 (3).	Jurkat (0.015), HepG2 (0.023), Sawano (0.041), MOLT-4 (0.044), HL-60 (0.066).	HL-60 (1.25), HepG2 (1.86).	CACO-2 (0.12), Sawano (0.24), MOLT-4 (0.47), TCS (0.71), HL-60 (0.72), HepG2 (1.9)	MOLT-4 (0.075), HepG2 (0.049), TCS (0.046), Hela (0.80), COS7 (0.045), Vero (0.050), Sawano (0.065)	HeoG2 (2.3). Hela (7.2)
Susceptible normal cell line (EC50µg/mL)	low to modest activity against lung cells. Less than 10 µg/mL towards UtSMC cells	T cells (0.148), Less than 10 µg/mL towards Hepatocytes	Less than 10 µg/mL towards T cells Less than 10 µg/mL towards Hepatocytes.	Substantial activity against T cells of native and recombinant PS-4. Recombinant PS-4 also cytotoxic to UtSMC.	Less than 10 µg/mL towards human embryonic lung fibroblast, uterus, and hepatocyte	Unknown

Table 2: Parasporin groups and their characteristics.
Human cell line origins: Jurkat (leukemic T-cell), HepG2 (liver cancer), Sawano (cervical cancer resistant to cisplatin treatment), MOLT-4 (acute lymphoblasic leukaemia), HL-60 (myeloid- granulocyte precursor in bone marrow- leukaemia), CACO-2 (epithelial colorectal adenocarcinoma cells), UtSMC (normal uterus smooth muscle cells), TCS (uterus cervix cancer). Adapted from Kitada *et al.* 2005 and Domanska, 2016 (Akiba *et al.*, 2009; Kitada *et al.*, 2006; Ohba *et al.*, 2009; Saitoh *et al.*, 2006; Yamashita *et al.*, 2005; Mizuki *et al.*, 1999; Mizuki *et al.*, 2000; Kitada *et al.*, 2005; Gonzalez *et al.*, 2011; Ammons *et al.*, 2016; Ekino *et al.*, 2014)

1.3 Mode of action of *Bt* toxins

The vast number of Cry toxins, the type, number of target cells as well as receptors has led to a number of theories on how Cry toxins are able to target, interact and kill particular organisms or cell lines.

Research has attempted to uncover the mode of action of *Bt* toxins and there is currently more than one theory on how the toxins are able to be toxic. Initially the general consensus accepted the model proposed by Knowles and Ellar. It was founded on studies carried out on 3-domain Cry toxins and their effect on lepidopteran larvae. They postulated that pathogenic bacteria usually exert cytolysis through pore formation and proposed the colloid-osmotic model (Knowles and Ellar, 1987).

In this model the insect must first ingest the *Bt* crystal which is then solubilised in the alkaline environment of the insect's digestive system. A study by Angus, (1954) highlighted the need for alkaline conditions as pre-requisite for solubilisation (Angus, 1954). The findings were later supported by observations made by Du *et al.* (1994) where only insects with digestive juice capable of crystal solubilisation were susceptible to Cry toxins (Du *et al.,* 1994; Jurat-Fuentes and Crickmore, 2017).

 Once solubilised the protoxin is cleaved by digestive proteases and results in protease resistant cores. This activated form binds to and interacts with specific receptors in the

gut border membrane. *In vitro* studies by Wolfersberger *et al*. (1984) demonstrated the correlation between binding and toxicity by using brush boarder membrane vesicles (BBMV). Early studies of Cry1Ba and BBMV from the susceptible *Pieris brassicae* found that Cry1Ba would specifically bind to insect BBMV and would not bind to rat intestine BBMV (Wolfersberger, 1984; Wolfersberger *et al.,* 1987). Post toxin-receptor interaction the toxin undergoes oligomerization.

A study by Güereca and Bravo, (1999) demonstrated that oligomerization of monomeric Cry toxins requires interaction with cell membrane (Güereca and Bravo, 1999). The events that follow have divided scientists and is the subject of a number of theories. In the colloid osmotic model, the toxin inserts itself into the membrane and as a consequence causes nonspecific pores in the membrane of the target cell. The presence of the pore induces changes in membrane potential, ion balance, influx of water, cell swelling, colloid-osmotic lysis of gut cells and eventual larval death (Knowles and Dow, 1993, Knowles and Ellar, 1987). There are several mode of action models that have since taken into consideration newly found data on receptors toxin complexes as well as the revealed crystal structure of Cry toxins.

1.3.1 The Bravo model

The Bravo model is founded on the osmotic -lysis model by Knowles and Ellar, (1987). Bravo *et al.* (2004) investigated the binding of Cry1Ab to two different receptors, aminopeptidase N (APN) and cadherin-like protein (Bt-R1 in *Manduca sexta*).

The model suggests that Cry1A toxin binds sequentially to these receptors in order to cause toxicity. It details the different monomeric changes that the Cry1A protein undergoes in order to cause toxicity.

The model addresses the first monomeric change where crystal Cry1A is solubilised to a 130 KDa protoxin. It is followed by a second monomeric change when it is activated and cleaved into a ˜60 KDa protease resistant protein. The third monomeric change takes place after active Cry1A binds to Bt-R1 receptors and then to APN receptors on the membrane of midgut cells of insects to create a toxin-receptor complex. The fourth monomeric change takes place as the toxin forms a pre-pore oligomeric structure that inserts itself into the membrane and subsequently creates lytic pores (Bravo *et al.,* 2004).

This mode of action is based on findings from the co-immunoprecipitation investigation where it was demonstrated that Cry1Ab binds first to Bt-R1 and then APN. As a consequence of this binding, membrane-associated proteases cleave the α-helix 1 of domain I and expose a hydrophobic surface that subsequently leads to the oligomerization of Cry1Ab into pre-pore structures.

 Experimental data by Soberón *et al*. (2007) demonstrated the formation of oligomeric structures by Cry1Ab and Cry1Ac mutant which lack the α-helix region. These mutants

exhibited *in vivo* toxicity in resistant insects with a mutated cadherin gene. Furthermore,

the oligomeric state of Cry1Ab-Bt-R1 complex greatly increased binding affinity to APN

which guides oligomers into the lipid rafts and enables insertion and the formation of

tetrameric pores (Bravo *et al.,* 2004; Soberón *et al.,* 2007; Soberon *et al.,* 2009; Bravo

and Soberón, 2008). Figure 8 summarises and illustrates the pore forming mode of

action of the Bravo model
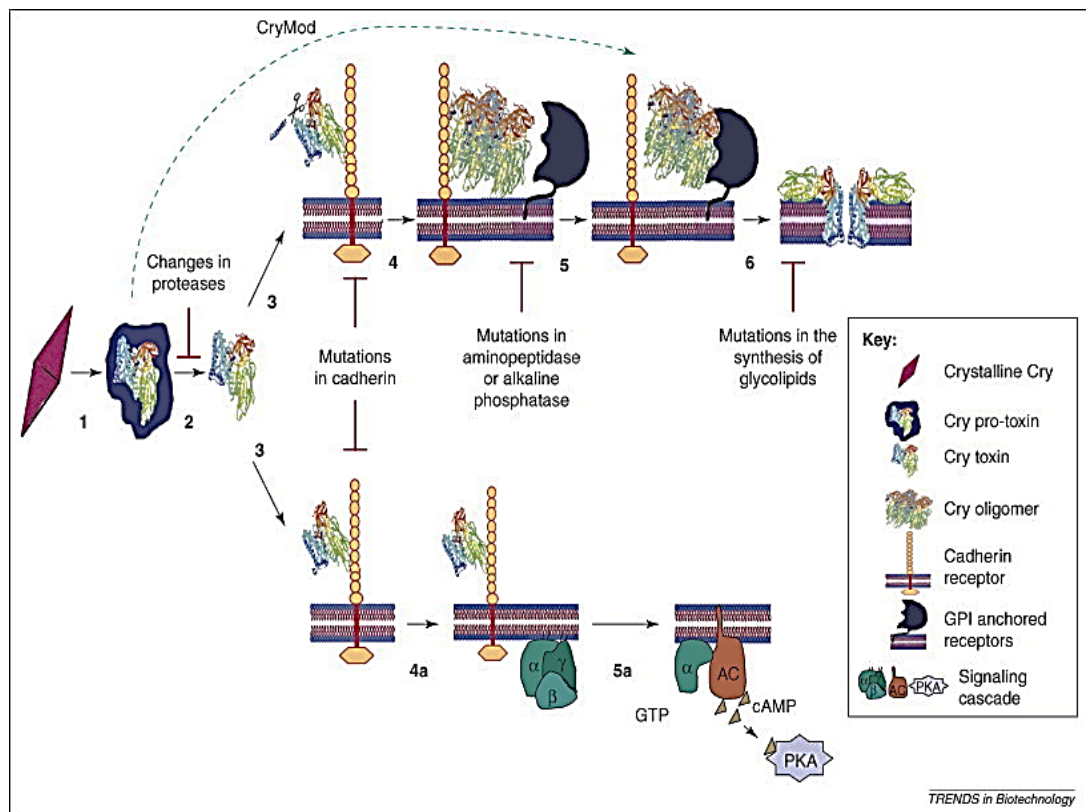
Figure 8 The two models for the mode of action of *Bt* Cry toxins.
Cry crystal is solubilised and activated by proteases. The active toxin binds to receptor in steps 1-3. The top model 4-6 illustrates the Bravo model. Binding induces the cleavage of a-helix. The monomers undergo oligomerization and bind to GPI-anchored receptor (Bt-R1). It induces membrane insertion, and pore formation. The lower diagram 4a-5a illustrated the signalling model by Zhang *et al.* (2005) It suggests a mode of action by an adenylyl cyclase/ protein kinase A signalling pathway. Image taken from Bravo *et al.* (2008).

1.3.2 The Ping Pong model

Similar to the Bravo model, and other mode of action models, the ping pong model agrees that Cry crystals must first undergo solubilisation, followed by protease treatment where the protoxin is cleaved resulting in a protease resistant core.

In the case of Cry1Ab the active ˜60 KDa core encompasses the typical Cry 3-domain structure (Pardo-López *et al.,* 2013). Pacheco *et al*. (2009) proposed the ping pong model where loop 3 of domain II Cry1Ab differentially participates in the binding of both Bt-R1 and APN receptors depending on the oligomeric state of the toxin.

Data from toxin overlay binding assay confirmed that amino acids from loop 3 in domain II interact with both receptors and that multi binding interactions take place which involves both receptors. ELISA binding assays suggested that loop 3 first binds to the high abundance low affinity APN receptor site followed by binding to the high affinity low abundance Bt-R1 receptor also known as a cadherin receptor (Pardo-López *et al.,* 2013).

Previous research in the cadherin interaction with loops of domain II revealed the involvement of loops α8, 2, and 3 (Gomez *et al.,* 2002; Gómez *et al.,* 2003; Gomez *et al.,* 2006). Loop2 was found to act as the cognate binding epitope to the CR7 region of cadherin, in addition, loop α8 and 2 also interacted with region CR11 of cadherin. Loop 3 was thought to bind to the CR12 region of cadherin receptor (Gomez *et al.,* 2006;

Gomez *et al.,* 2003). Pacheco *et al.* (2013) confirmed loop 3 binding to CR12 region after monomeric mutants of Cry1Ab loop 3 were affected in their ability to bind to CR12 fragments of Bt-R1 but did affect interactions with APN receptor. Pacheco *et al.* (2013) also suggested that binding to high affinity Bt-R1 requires the participation of other domain II regions, e.g. loops 2 and α8. Once bound, the Bt-R1 loop 3 complex induced the cleavage of α-helix 1 to form an oligomer that increased binding affinity to APN via β16-22 of domain III. Whilst bound to Bt-R1 via loop 3 a Bt-R1-Cry1Ab-APN complex is created. However, it is APN that the facilitates oligomer insertion into the membrane. Figure 9 summarises the 'ping pong' model as proposed for Cry1Ab mode of action.
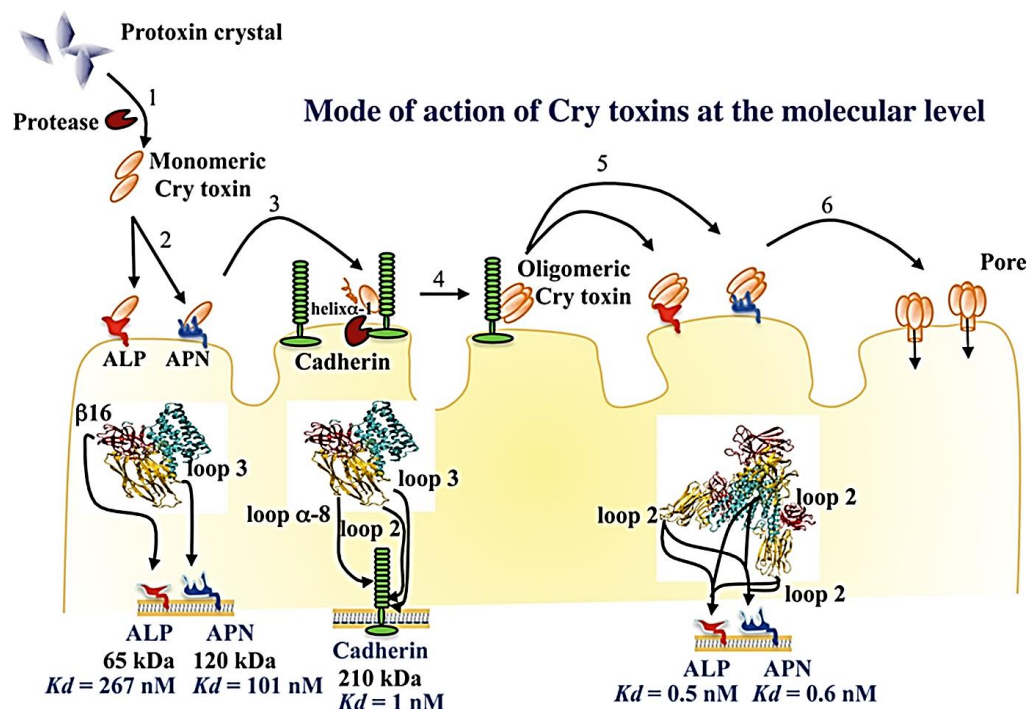
Figure 9 The 'ping pong' model for Cry1Ab mode of action.
Cry1Ab protoxin is activated with proteases. The monomeric toxin binds first to low affinity ALP or APN receptors via loop 3. This is followed by further binding to high affinity cadherin or Bt-R1. The N-terminal is cleaved and toxin monomers oligomerize into pre-pore structures that bind APN or ALP to induce pore formation. Figure taken from Pardo-López *et al.* (2013).

91.3.3 The Zhang model

The Zhang model is an alternative model that is based on the adenylyl cyclase (AC) protein kinase A (PKA) signalling pathway (Zhang *et al.,* 2005). Zhang *et al.* (2005) argue that the toxic events take place as a consequence of the signalling pathways and not a result of pore formation.

Once the crystal is solubilised and activated, the monomeric Cry1Ab binds specifically to Bt-R1 receptors. This would induce a signalling pathway where oligomers do not form lytic pores (Zhang *et al.,* 2005). The model is summarised in the lower part of diagram 7.

The research was carried out using undifferentiated ovarian cells of the cabbage looper *Trichoplusiani* which does not express cadherin Bt-R1 receptors. The cell line known as S5 was engineered to express Bt-R1 driven by an insect action promoter in a nonlytic insect expression vector. The receptor was expressed as a membrane protein and localised to the cytoplasmic membrane of cell as confirmed by immunofluorescent staining using anti-BT-R1 antibody. Western blot experiments concluded that the monomeric form of the toxin was taken into the membranes expressing Bt-R1 receptors.

The uptake of monomeric toxin was fast and steady in contrast to oligomeric form of Cry1Ab which was not detected by either Bt-R1 expressing cells or other non-susceptible cells. In a dose dependent experiment oligomeric toxin was detected 15min post incubation in Bt-R1 expressing S5 cells and did not exert a toxic effect.

Zhang *et al.* (2005) concluded that it is the monomeric from that binds to Bt-R1 to exert toxicity in susceptible cells (Zhang *et al.,* 2005). Zhang *et al.* (2006) proposed that this binding triggers a G protein-coupled receptor (GPCR) which in turn activates the AC/PKA pathway. *extra*cellular ligand activation of GPCR causes a conformational change on the intercellular region of the membrane. The $G_{\alpha s}$ subunit switches GDP to GTP which in turn disconnects from the complex and binds instead to AC. AC induces the rise of a second messenger cAMP in the cytoplasm of susceptible cell that triggers PKA.

There was evidence of raised cAMP and PKA levels in susceptible cells. A decrease in toxicity levels was observed with the application of $G_{\alpha s}$ and AC inhibitors. Toxicity was completely abolished with the addition of EDTA pre-incubation with toxin. Loss of toxicity was not due to loss of binding as observed from western blots experiments but as a consequence of Mg^{2+} depletion. Toxicity was re-established with the addition of magnesium salt. Zhang *et al*. (2006) claimed that Mg^{2+} is critical to events that follow binding of monomer toxins to membranes of susceptible cells and that receptor-toxin interaction and pore formation are not enough to explain Cry1Ab mode of action (Zhang *et al.,* 2006).

A summary of the Zhang model is illustrated in figure 10. This claim was supported by research carried out in modified Cry1 toxins that lacked the N-terminal ending including helix α-1, where cadherin interactions were skipped and did not take place. Despite this resistant insects with a defective cadherin gene were killed (Soberón *et al.,* 2007).

Other studies have criticized the experimental data on which the Zhang model is based on. They claim that the S5 cell line does not represent the true nature of a susceptible insect's midgut. Furthermore, EDTA amount used by Zhang *et al*. (2006) would not chelate all $MgCl_2$ that would be present in an insect medium where low pH level would affect ion chelating by EDTA and EGTA. (Vachon *et al.,* 2012; Soberón *et al.,* 2007).
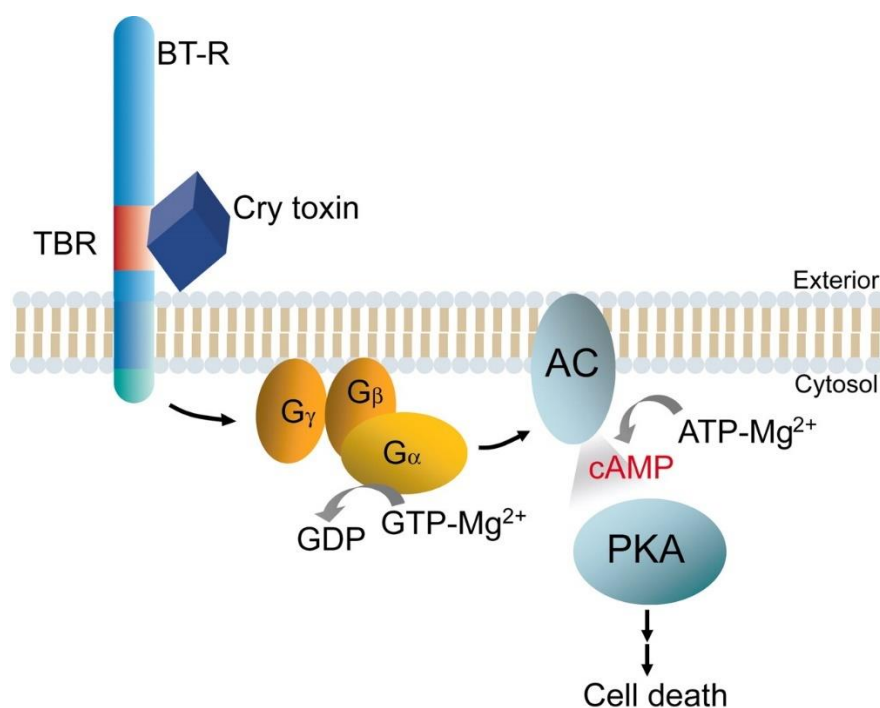
Figure 10 The mode of action of *Bt* toxin according to the Zhang model.
It implicates AC/PKA signalling pathway. It states that monomeric form of the toxin is required in binding to cadherin receptor. The binding induces coupled G protein which in turn activates adenylyl cyclase and a raise in cytoplasmic cAMP. cAMP levels activate kinase A which phosphorylates target proteins resulting the disruption of channels and eventual cell death. Figure taken from Zhang *et al*. (2006)

1.3.4 The Jurat-Fuentes interpretation

The Jurat-Fuente, (2006) interpretation combines both the Bravo and Zhang models. It suggests that both pore formation and signalling pathway play a role in mode of action of Bt toxins and the presence of one does not eliminate the other.

Jurat-Fuentes and Adang, based their alternative interpretation on Cry1Ac research carried out in the susceptible *Heliothis virescens* larvae. Previous research identified a number of receptors for Cry1Ac *in H.virescens* (Jurat-Fuentes and Adang, 2006, Jurat-Fuentes and Adang, 2004). These included HevCaLP, HvALP, actin receptors and intracellular phosphatases. A gene knock out of the cadherin like HevCaLP receptor was found to cause resistance in certain insect strains and was implicated in the toxin mode of action (Gahan *et al.,* 2001).

HvALP is a GPI-anchored alkaline phosphatase receptor which was found at low levels in resistant *H.virescens* strains that exhibited reduced alkaline activity phosphatase in brush boarder membrane proteins (Jurat-Fuentes and Adang, 2004). Additional receptors were also implicated, the discovery of more receptors such as actin in BBMV proteome of *M .sexta* and *H. virescens* (Krishnamoorthy *et al.,* 2007; McNall and Adang, 2003) and intracellular phosphatases (Jurat-Fuentes and Adang, 2006) has led to the proposal of two modes of action by Cry toxins.

In the first model, Cry1Ac binds to cadherin like HevCaLP, the binding causes a conformational change. Post oligomerization, the toxin binds to the HvALP receptor which subsequently results in pore formation in lipid rafts of the membrane. The presence of the pore creates an osmotic shock that eventually results in the cell death.

The second model speculates that toxin binding to HevCaLP does not just induce pore formation but that it also activated signalling pathway regulated by intracellular phosphatases. This in turn activated the apoptotic pathway (Jurat-Fuentes and Adang, 2006). Lilien and Balsamo. (2005) implicated the cytoskeleton protein actin in cadherin interactions. Jurat-Fuentas. (2007) suggest that it binds to cadherin via its cytosolic domain which is aided by tyrosine phosphatases. It is also possible that actin binds to oligomerized toxin to induce the signalling pathway (Jurat-Fuentes and Adang, 2007., Lilien and Balsamo., 2005). The mode of action of Cry1Ac in *H. virescens* as interpreted by Jurat-Fuentas. (2007) is summarised in figure 11.
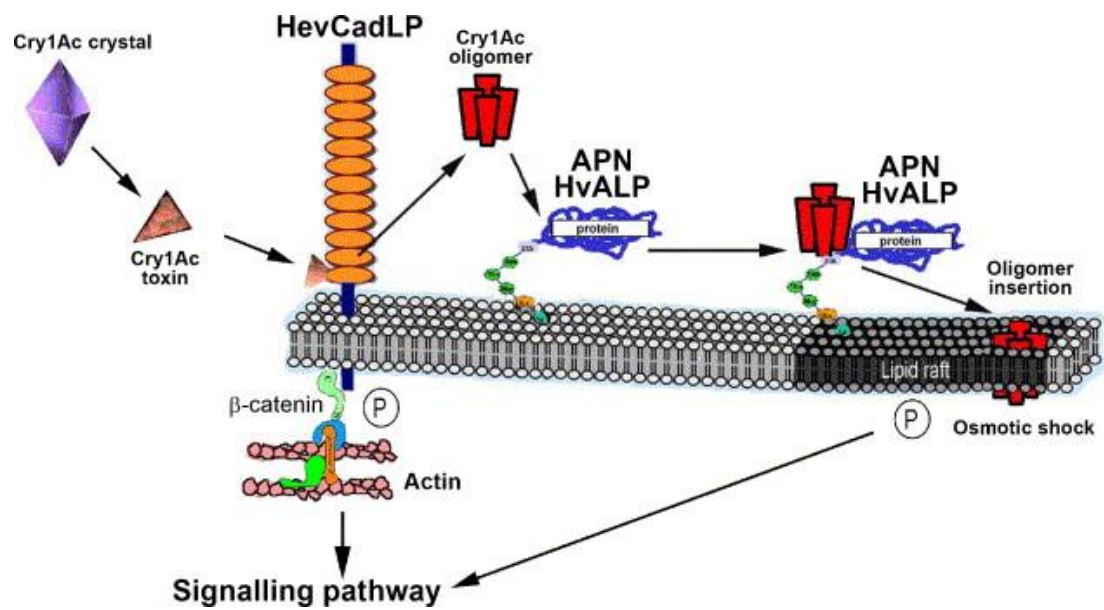
Figure 11 The proposed mode of action of Cry1Ac in *H. virescens.*
Once the crystal is solubilised and activated through protease, it binds to HevCaLP. The oligomerized toxins binds to second receptor HvALP the leads to membrane insertion and subsequent pore formation and cell death. Alternatively, the second model proposes that toxin bound to cadherin induces interactions between actin and cadherin that activate a signalling pathway. Taken from Jurat-Fuentes and Adang *et al*. (2006)

1.4 Specificity of *Bt* toxins

The topic of specificity in *Bt* toxins goes hand in hand with *Bt* toxin mode of action. A number of studies have implicated the structure of *Bt* toxins with their specificity to target particular insects (de Maagd*et al.,* 2003; Jez, 2017; Moar *et al.,* 2017).

Adang *et al*. (2014) summarised the mode of action of 3-domain toxins where domain II and III interact with the midgut epithelial cell of insects whilst domain I was responsible for membrane insertion and subsequent pore formation (Adang *et al.,* 2014). However in a recent review it was argued that there is not sufficient evidence which directly links

toxin structure and its specificity towards a target (Jurat-Fuentes and Crickmore, 2017). A review by Palma *et al*. (2014) provided support to this notion where it stated that that Cry toxins can exhibit toxicity against a range of taxonomically diverse insects, and Cry toxins with a diversity in domain II and III can still target the same insect (Palma *et al.*, 2014; Palma and Berry, 2016). Jurat-Fuentes and Crickmore (2017) , argue that toxin structure alone is a poor indicator of specificity and that host factors may also play a role in determining specificity (Jurat-Fuentes and Crickmore, 2017). Previous studies on insects that demonstrate toxin resistance concluded that they are able to do this as a result of alternations to the receptor binding site on host cell, and emphasising the specificity of toxins to target insect (Bravo and Soberón, 2008).

Jurat-Fuentes and Crickmore, (2017) further speculate that parasporins or Cry toxin which exhibit toxicity to certain cancer cell lines may have insect targets that are yet to be discovered (Jurat-Fuentes and Crickmore, 2017; Krishnan *et al.,* 2017). The authors defined specificity 'as the condition of Cry proteins being toxic to a particular insect' and propose that there are seven levels of specificity determinates (Jurat-Fuentes and Crickmore, 2017). Figure 12 summaries the seven levels in the mode of action of Cry insecticidal proteins that determine toxin specificity. Level I address an insect's ability to access or come across a *Bt* crystal as the first level of determining specificity. The crystal form of the toxin limits the range of insects that can access it e.g. sap feeding hemipterans. *Bt* is also limited by the type of environments that it can successful inhabit.

For example, *Bt* do not usually colonize plant leave surfaces and insects that live there are unlikely to encounter *Bt* or its crystal.

Parasporins which exhibit specificity to certain cancer cell lines are not thought to have evolved this specificity. Instead Jurat-Fuentes and Crickmore, (2017) suggest that parasporins may have undiscovered insect targets. This would not be unusual as previous Cry toxins have shown specificity to more one species in an insect order as is the case for 6 of 68 Cry toxin families (Jurat-Fuentes and Crickmore, 2017).

Level II of specificity is the ability of a host to solubilise the *Bt* crystal. Early studies on crystal solubilisation confirmed that *Bt* crystals require alkaline pH conditions and insects which lack such a digestive environment were unaffected by *Bt*. Further investigation confirmed that previously non-susceptible insects were affected by pre-solubilised *Bt* crystals (Angus, 1954; Du and Nickerson, 1996; Du *et al.,* 1994).

Specificity level III addresses toxin processing and the stability of protease cleaved *Bt* toxins. Once solubilised the protoxin is available to digestive juices in the insect gut. There digestive proteases cleaved the protein until a protease resistant protein(s) core remains. The speed and efficacy of processing determines a toxins level of toxicity towards an insect.

Specificity level IV requires that the active toxin remain available and stable in insect gut juices. Although the active protease core is not further processed by a given protease it is not protected from further processing by others e.g. elastase (*Choristoneura fumiferana*) or hexamerins (*H. armigera*) have both reduced toxicity in susceptible Cry1A insects by sequestering the active toxin.

Specificity level V address the task of crossing the peritrophic matrix which protects epithelium cells of the insect gut. The layer is made of chitin and glycosylated proteins. Some Cry toxin domains recognise and interact with glycosylic residues which stall and hinder the intoxication process. It has also been observed that the type of glycosylated protein and their arrangement differ in insect's species, making this a determining factor for the specificity to a particular target insect.

The specificity level VI examines a toxin's ability to recognise, bind and interact with insect cell receptor (Jurat-Fuentes and Crickmore, 2017). Previous studies in specificity have addressed receptor-toxin interaction as the Key determining factor of a toxin's specificity towards an insect. Studies have identified Cry toxin receptors that include protein and glycolipids (de Maagd *et al.,* 2001; Gomez *et al.,* 2007; Vachon *et al.,* 2012; Bravo *et al.,* 2007). The specificity of parasporins to certain cancer cell line is not well understood, it is thought that parasporins recognise carbohydrate structures shared with the susceptible cancer cell line but which are absent in normal healthy cells (Jurat-Fuentes and Crickmore, 2017). Studies carried out in Cry1Ab have shown that this

insecticidal Cry toxin is cytocidal to human embryonic kidney cell line 293. The authors suggest that it is not only parasporins that can interact with mammalian cell lines, but new specificity is emerging on typical insecticidal Cry toxins as research widens the range and type of cells being tested against Cry toxins (Mesnage *et al.,* 2013).

The final specificity level VII addresses events that follow toxin-receptor binding. The ability to bind does not necessarily guarantee pore formation and to induce a toxic effect. Studies have shown high binding affinity in low susceptible insects. Different cells can have different responses to membrane insertion that do not always end in lysis of cell. In addition to this cells can activate different intracellular signalling pathways with different defensive response to toxicity (Jurat-Fuentes and Crickmore, 2017).
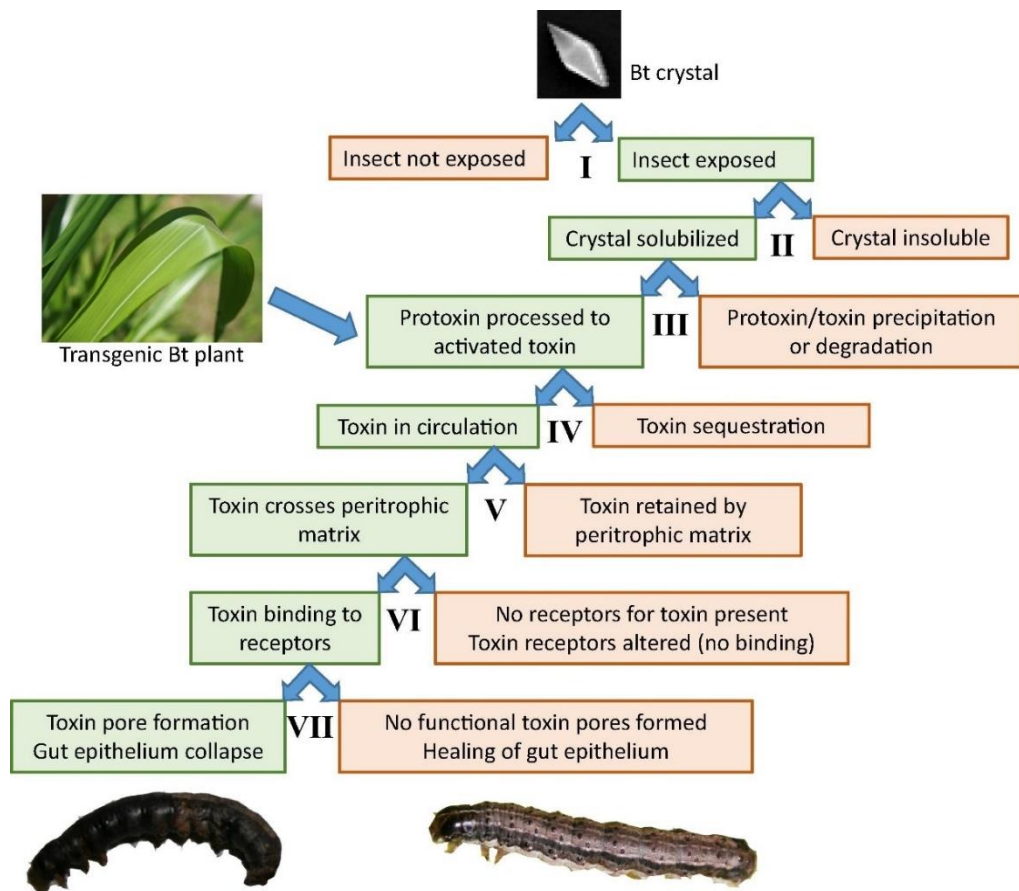
Figure 12 The flow chart diagram of the seven steps of the mode of action of Cry toxins and the determining specificity factors.
Figure depicts factors that play a role in determining host specificity. Transgenic *Bt* crops produce soluble Cry toxins and thus are excluded from the first two specificity determinants. Taken from Jurat-Fuentes and Crickmore, (2017) (Jurat-Fuentes and Crickmore, 2017).

1.4.1 Domain II loops in Cry toxins

A number of studies have implicated domain II loops in specificity and initial binding of

Cry toxins (Dean *et al.,* 1996; Pacheco *et al.,* 2009; Abdul-Rauf and Ellar, 1999; Pardo-

Lopez *et al.,* 2009; Abdullah and Dean, 2004; Lu *et al.,* 1994).

Pacheco *et al.* (2009) demonstrated that mutations in domain II loops affect binding to the cadherin (Bt -R1) and aminopeptidase-N (APN) receptors in the lepidopteran *Manduca sexta*. The study suggested that both receptors interact differently to different oligomeric states of the Cry1Ab toxin. Domain II exposed loop regions were probed and mutations in loop 3 were found to severely affected its insecticidal activity in both monomeric and oligomeric states of the activated toxin. Pacheco *et al.* (2009) proposed a ping pong binding mechanism to explain the toxin activity and its dependency on loop 3 of domain II.

Toxin overlay binding analysis of loop 3 regions of Cry1Ab indicated that it has hydrophobic profile. Furthermore, single double and combination substitution in this region revealed that this region is involved in interaction with both Bt -R1 and APN receptors as binding to both receptors was significantly reduced in mutants. More analysis suggested that the Cry1Ab mode of action may involve multiple binding interactions with both receptor molecules. Both receptors are present at different concentrations with APN at a higher concentration in *M. sexta* mid gut.

Initially a 65 KDa proteolytic active monomer binds to APN through loop 3 of domain II. This bound monomer goes on to bind to high affinity Bt -R1 receptor through loops 2, α-8, and 3 of domain II creating an oligomer. Mutants of loop 3 were noted to significantly affect oligomeric binding to Bt -R1. The resulting oligomeric structure has binding affinity to APN through domain III regions but continues to be bound to Bt -R1 through loop 3.

The interactions result in a Bt -R1 -Cry1Ab-APN complex. The specific role of this complex is unclear, but the study observed that loop 3 mutants affect toxicity of Cry1Ab oligomers and thus implicated the complex in toxicity. It further suggested that loop 3 can induce structural changes prompted by APN and it is involved in membrane insertion of Cry1Ab oligomer.

Previous research on loop 3 have also suggested that it binds to CR12 region of cadherin receptors in *H.virescens* and *B. mori* (Xie *et al.,* 2005; Atsumi *et al.,* 2008; Pacheco *et al.,* 2009). Rajamohan *et al.* (1996) carried out alanine substitution on Cry1Ab, which is toxic to M. *sexta* and *H. virescens*. Two mutants F440A, and G439A were noted to have reduced toxicity and binding analysis on brush boarder membrane vesicles from insect gut membranes implied that this reduction in toxicity is due to loss in the initial binding by as much as 3.5 times less than wild type Cry1Ab.

Alanine substitutions made to hydrophobic residues [440] AAGA[443] in loop 3 of Cry1Aa also resulted in reduced toxicity to *B. mori* and *M. sexta* due to a decrease in initial binding to vesicles from insect midgut. The study proposed that residues of loop 3 of both Cry1Ab and Cry1Aa ascertain hydrophobic interactions with receptor molecules; and it is the hydrophobic nature of these residues that affected the initial binding to receptor. To investigate further, an alanine cassette was created in the loop 3 region of Cry1Aa and resulted in the deletion of hydrophobic residues in the loop 3 region. This lead to a

reduction of binding affinity (K_{com}) by 15 times to *B. mori* and 9 times to *M. sexta*, implicating the loop 3 region in initial binding to receptor (Rajamohan *et al.,* 1996c).

Alanine substitutions of the predicted loop 3 of Cry1C residues indicated that single point mutations S438F and G439A eliminated most or all toxicity and binding to *S. littoralis* and *A. aegypti*. The mutants expressed a stable proteolytic resistant protein similar to that produced by insect gut juices, and thus the loss of toxicity was not thought to be due to protein mis-folding (Abdul-Rauf and Ellar, 1999). Cry4Aa is toxic to *B. mori* and *M. sexta*. Studies which investigated the effects of alanine substitutions in loop 3 in Cry41Aa concluded that mutants T512A and Y513A have a reduced toxicity to susceptible insects (Howlader *et al.,* 2010; Howlader *et al.,* 2009).

In a study carried out in Cry1Aa to investigate the hydrophobic nature of loops in domain II, it was questioned whether hydrophobic profiles of loop residues were conserved in Cry toxins or necessary for initial binding to the Bt -R1 cadherin receptor. Seven mutants of loop 3 of Cry1Aa were created, and none exhibited a reduction in receptor-binding affinity. Five recombinants out of the seven produced 75% mortality rate in third instar *B. mori* within 72 h of ingesting recombinant toxins. One recombinant (^{439}EPPA^{442}) caused a 35% mortality rate, whilst recombinant ^{439}TLRT^{442} was completely non-toxic with zero morality at both 48 and 72 h post ingestion of the recombinant. The study concluded that the critical residues are not required for binding. However, the study emphasised the natural flexibility of loops of Cry1Aa and the possibility that mutants

may change conformation to present hydrophobic profiles to the cadherin receptor and thus averted a reduction in receptor binding affinity (Fujii *et al.,* 2013).

Fujii *et al.* (2013) speculated that the natural flexibility of loop allowed for multipoint attachments by exposing hydrophobic profiles to receptors, without the need for specially conserved amino acid sequence in the loop region. Furthermore, the study suggested that larger conformational flexibility of loops involved in binding allowed for a larger available fraction of loop conformation compatible with toxin-receptor complex formation. This may explain their observations whereby some of loop 3 mutants exhibited higher binding efficiency by up to 42 times.

2.0 Objectives

A number of studies have suggested that the loops of domain II are involved in initial binding and specificity of Cry toxins. This study aims to identify domain II loop regions of Cry41Aa in order to investigate the specificity that it has towards HepG2 cancer cell lines.

Cry41Aa is one of the least studied parasporins. Little is known about what gives rise to its specificity towards some mammalian cancer cell lines. This study will apply functional and structural approaches that have been typically used to study insecticidal Cry toxins, in order to unravel Cry41Aa specificity. Bioinformatic tools and findings from previous research on insecticidal Cry toxins will server to identify and narrow down Cry41Aa gene regions corresponding to domain II loops regions. Once identified the domain II loops will undergo the following mutagenesis approaches: (I) a loop exchange between cytocidal Cry41Aa and an insecticidal Cry toxin. (II) the production of hybrids as a result of domain swapping between Cry41Aa and other insecticidal Cry toxins. (III) the creation of Cry41Aa loop mutants via residue substitutions. The above approaches will contribute to building a picture of Cry41Aa specificity and identify key gene regions or residues that are significant to its specificity.

The long-term project outcomes of this study aim to contribute to the current discussion on the safe application of Cry toxins as biopesticides and the discovery that Cry toxins can and do interact with vertebrate cell lines. Findings from this study may reveal alternative or unknown HepG2 receptors targeted by Cry41Aa. Any further information

learnt on the specificity of Cry41Aa can be used to assess its potential contribution to the pharmaceutical industry and drug design in the fight against cancer.

The long-term project outcomes of this study aim to contributes to the current discussion on the safe application of Cry toxins as biopesticides and the discovery that Cry toxins can and do interact with vertebrate cell lines. Findings from this study may reveal alternative or unknown HepG2 receptors targeted by Cry41Aa. Any further information learnt on the specificity of Cry41Aa can be used to assess its potential contribution to the pharmaceutical industry and drug design in the fight against cancer.

3.0 Material and Methods

3.1 Materials

3.1.1 Bacteria Strains and Plasmids

Genomic DNA of A1462 *Bt* strain was used as a DNA template to amplify Cry41A operon (Yamashita *et al*., 2005). DNA inserts were ligated into pBluescript or *E.coli*/ *Bt* shuttle vector pSVP27A (Crickmore *et al*., 1994). *B. thuringiensis/Escherichia coli* shuttle vector (*Bt*-shuttle vector) pSVP27A is characterised by its ability to replicate in both *E.coli* and *Bt* (Liu *et al*., 2009). It has a promoter and a selective marker DNA sequence that can function in *Bt*.

In this study the *Bt*-shuttle vector pSVP27A was modified from vector pSV2 which is partly derived from pBR322 vector (Crickmore and Ellar, 1992; Crickmore *et al*., 1994). *E. coli JM109* strain was transformed with plasmid to establish constructs. Other plasmids include pGEM which is high-copy-number vectors containing T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the α-peptide coding region of the enzyme β-galactosidase. A list of host bacterial strains is presented in table 3. Table 4 lists the buffers used in this study. Table 5 lists the regents and kits employed in this study.

Bacteria Host List

Bacterial strain	Description	Bacterial host
JM109	high efficiency transformation strain used for routine cloning	*E. coli*
GM2163	dam⁻/dcm⁻ chloramphenicol resistant strain used to obtain non-methylated dna for transformation of *Bt4D7*	*E. coli*
Bt4D7	crystal minus derivative of bt subspecies *kurstaki,* obtained from the bacillus genetic stock centre	*Bt*

Table 3 lists the bacterial strains used in this stud

3.1.2 Buffers and Solutions

List of Buffers and Solutions

Buffers	Composition
TBE (5x)	108 g of tris, 55 g of boric acid, 40 mL of 0.5 M EDTA, 2 l of dH2O, pH 8.0.
RGB	18.18 g tris, 0.4 g SDS, 100 mL of dH2O, pH 8.8.
SGB	6.06 g tris, 0.4 g SDS, 100 mL of dH2O, pH 6.8
PROTEIN GEL SAMPLE LOADING BUFFER (2x)	2 g SDS, 6 mg EDTA, 20 mg Bromophenol Blue, 5 mL of RGB, 50 mL glycerol, 100 mL of dH2O.
SDS RUNNING BUFFER (10x)	7.6g tris-HCl, 36g glycine, 2.5g SDS, 250 mL of dH2O
COOMASSIE BLUE	methanol, dH2O, acetic acid (10:9:1 v/v/v), Brilliant Blue R-250 (0.25%, w/v)
Destain	methanol, dH2O, acetic acid (10:9:1, v/v/v).
PBS (10x)	80 g of 1.37 M NaCl, 2 g of 27 mM KCl, 14.4 g of 100 mM Na2HPO4, 2.4 g of 18 mM KH2PO4, 1 l of dH2O, pH 7.4
RSS solution	6.5 mL dH$_2$O, 100 μL of 500 mM EDTA, 1 g 30 % sucrose, 150 μL of 10 % SDS, 2.5 mL of 0.5 M NaOH, 600 μL of 1.4 M KCL, Bromophenol blue (BPB) dye
50 mM Sodium Carbonate pH 10.9	0.26 g of Na$_2$CO$_3$ solute in 50 mL of dH$_2$O.
50 mM Sodium Hydrogen Carbonate pH 8.3	0.21 g of NaHCO$_3$ solute in 50 mL of dH$_2$O.
CAPS (10mM)	2.21g CAPS, 1l dH2O pH10.5
CAPS (10mM) + NaCl (1M)	2.21g CAPS, 1M NaCl, 1l dH2O pH10.5
NP-40	150 mM NaCl, 1.0% NP-40 or Triton X-100, 50 mM tris, pH 8.0.
RIPA	150 mM NaCl, 1.0% NP-40 or Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM tris, pH 8.0.
DRY BLOT	39 M glycine, 48 mM tris, 0.0375% SDS, 20% methanol.
TRIS-HCL	20 mM Tris HCl, dH2O pH 7.5
ECL SOLUTION	10 mL of 100 mM tris pH 8.5, 3 μl of H2O2, 25 μl of 14.7 mg/mL p-coumaric acid, 50 μl of 88.6 mg/mL luminol.
NP-40	150 mM NaCl, 1.0% NP-40 or Triton X-100, 50 mM tris, pH 8.0.
PATCH CLAMP NACL BUFFER	140 mM NaCl, 5 mM KCl, 1.1 mM MgCl2, 1.1 mM CaCl2, 10 mM Hepes (pH 7.4).
PATCH CLAMP KCL BUFFER	140 mM KCl, 1.1 mM MgCl2, 0.1 mM EGTA, 10 mM Hepes (pH
PLB KCL BUFFER	150 mM KCl, 1 mM CaCl2, 10 mM Hepes, (pH 7.5).
BUFFER 1, BUFFER 2, BUFFER C	10 mM Tris-HCl , 2.5 mM MgCl2 , 0.5 mM CaCl2, (pH7)

Table 4 Lists buffers used in study

3.1.3 Reagents, kits, enzymes, plasticware

list of Reagents

Company name	Reagent/kits/enzyme
SIGMA-ALDRICH	β-mercaptoethanol, Bromophenol Blue, Brilliant Blue R-250, ammonium persulfate, TEMED, acrylamide/bis-acrylamide 30%, SDS, tris base, tris-HCl, CAPS, sodium carbonate, hydrogen peroxide, DMSO, chloramphenicol, etoposide, 5-Fluorouracil, S-methylmethanethiosulfonate (MMTS), EGTA, p-coumaric acid, luminol, sodium orthovanadate, sodium arsenite, proteinase K, lysozyme, formaldehyde solution, RNase ZAP, ethanol absolute for molecular biology. Vivaspin 20 centrifugal concentrator MWCO 100 kDa
BIOTUM	GelRed nucleic acid gel stain.
QIAGEN	QIAprep Spin Miniprep Kits, QIAquick PCR purification Kits. Exact buffer compositions are confidential, where possible detail has been provided. QG buffer: solubilization and binding buffer Chaotropic PB buffer: guanidine hydrochloride and isopropanol PE buffer: not available EB buffer: 10 mM Tris-Cl, pH 8.5 P1 buffer: resuspension buffer used when purifying plasmid DNA. P2 buffer: lysis buffer used when purifying plasmid DNA. Buffer N3: neutralization buffer used when purifying plasmid DNA. EL buffer: lysis buffer
MELFORD	trypsin, agarose low EEO, ampicillin
AANALAR BDH	glucose, NaCl, NaOH, Ponceau S, NP-40, TX-100, ethanol, sodium hydrogen carbonate.
FISHER SCIENTIFIC	Acetic Acid Glacial, FCS, 40 μm Nylon Mesh Cell Strainer, Nalgene 2.0 mL cryogenic vials, Trypan Blue HyClone (0.2 μm filtered)
THERMO FISHER SCIENTIFIC	glycine, methanol, 1-butanol, phosphate, acetic acid, trypsin (MS-grade), HRV 3C Protease, RNase inhibitor, High Capacity cDNA Reverse Transcription kit, Power Syber Green PCR master Mix, glycerol, glycine.
NEW ENGLAND BIOLABS	Pre-stained Protein Ladder (7-175, 11-245 and 10-230 kDa), 1 Kb DNA ladder, *Dpn*I, T4 DNA ligase
TOCRIS	BSA, PKi (14-22 amide, myristoylated), 8-bromo-cAMP
Roche	Complete Mini EDTA-free Protease Inhibitor Coctail Tablet
PROMEGA	HaeIII, buffer C CellTiter-Blue Cell Viability Assay, CellTox Green Cytotoxicity Assay,
GIBCO -LIFE TECHNOLOGIES	DMEM (high glucose), Advanced DMEM (high glucose), DMEM (high glucose, no phenol red), DMEM (low glucose), RPMI 1640, MEM, PSG, DPBS, DPBS with Ca2+ and Mg2+, trypsin/EDTA (0.05% trypsin and 0.53 mM EDTA), Trypan Blue (0.4%)
NUNCLON SURFACE	25, 75 and 175 cm2 flasks; 60 mm dishes; 6, 12, 24 and 96-well clear and black-walled clear flat bottom plates
CORNING	Poly-D-Lysine treated 96-well white and black wall clear bottom plates (3843) and cell scrapers (1.8 cm).

Table 5 lists the reagents, kits, enzymes and plasticware used in study.

3.1.3. Cell lines

HepG2, human hepatocyte carcinoma cell line was purchased from ECACC, Salisbury, UK. HeLa, lymphoblastoid (GM12878, IB4) and Burkitt's lymphoma (MUTU1, BL31) cells were gifts from Dr Michelle West (University of Sussex, UK). HL-60 cell line was a gift from Dr Helen Stewart (University of Sussex, UK). Another batch of HL-60 cells was purchased from ECACC, Salisbury, UK. Cell lines used in electrophysiology experiments were provided by Prof Jean-Louis Schwartz (University of Montreal, Canada).

3.1.4 Storage of biological material

Bacterial strains were storage in LB with 15% glycerol at - 80°C. Once defrosted bacteria were stored on agar plates at 4°C. Toxins were stored at - 20°C. Cancer cells were kept in liquid nitrogen at - 197°C, once defrosted for use they stored at - 80°C in complete medium with 10% glycerol.

3.2 Methods

The steps and experimental procedures taken to create *Cry41Aa* constructs are summarised in figure 13.

3.2.1 Agarose Gel Electrophoresis

All Agarose gels and DNA samples analysed by agarose gels in this study were prepared as followed unless otherwise indicated. 1% DNA agarose gel was made by weighting

0.3 g of agarose and added to 30 mL of TBE x1 and heated until all dissolved. A 1.5 μL of GEL RED dye was added and the solution was allowed to cool and solidify. A 1μL of purified product was added to 9μL of water, 2 μL of loading buffer was added to mixture and a total volume of 12 μL for each sample was loaded onto gel. A 1 Kb DNA ladder (New England Biolabs) was ran on gel at 120v in TBE buffer to make appropriate DNA band size comparisons. The bands of amplified products were visualised under ultraviolet light.

3.2.2 Amplification of Cry41Aa Operon and other DNAs via PCR

Genomic DNA of A1462 *Bt* strain was used as a template to amplify ORF2 and ORF3 (Krishnan, 2013). OFR2 of Cry41Aa was cloned into pBluescript SK2+ plasmid to create pBS41Aa. A BamHI restriction site at the 5' end and an XhoI site at the 3' end flanks the gene. The restriction enzymes allow for subcloning of ORF2 into pSVP2741Aa. The PCR amplification was carried out using high fidelity polymerase in the presence of dNTPs. The protocol for the PCR was as follows: initial denature was carried out at 94°C for 15s, the annealing temperature was 50°C for 30s, extension was carried out 68°C for 5 min with a final extension at 72°C for 7 min. This was repeated for 30 cycles. The PCR reaction was set up as follows: 0.5 μL of 100 μg/mL of each primer, 25 μL of high-fidelity PCR master mix (Roche), 0.5 μL of template DNA obtained from pBS41Aa plasmid, 2.5 μL of PCR nucleotide mix, and 23.5 μL of water. Amplified products then visualised using 1% agarose gel. All constructs were created using this protocol with occasional adjustments

to extension time and annealing temperatures. Primer are listed for each construct in the results chapter 5,6, and 7.

3.2.3 PCR Product gel purification

QIAquick PCR Purification Kit (QIAGEN) was used to purify PCR products according to manufactures recommendations. PCR products were run on agarose gel. The bands were visualised and excised. 600 μL of QG buffer was added and placed in a 60°C water bath until agarose gel solubilised. 200 μL of isopropanol was added and the solution was thoroughly mixed. 5x volume of chaotropic buffer PB is added to 1x volume of PCR product and mixed. QIAquick spin column is placed in 2 mL collection tube. The sample was added to column and spun for 1 min at 13,000rpm; this step binds DNA onto the membrane in the column. The flow through was discarded and column placed back in the same collection tube. A total of 750 μL of PE buffer was added to column and spun for 1 min. Second round of flow through is now discarded. Column is placed back on collection tube and further centrifuged for 1 min. Final round of flow through was discarded. The column is placed in a new 1.5 mL eppendorf tube. A total of 10 μL of EB buffer (10 mM Tris-Cl, pH 8.5) was added to centre of membrane and allowed to stand for 1min to maximise product recovery. Then it was centrifuged for 1 min. Each purified PCR product was collected and ran on 1% agarose gel.

3.2.4 DNA ligation

All purified linear DNAs were ligated overnight at room temperature. This will allow the linear DNA to become a circular one, and thus a plasmid that can be introduced into host bacteria. Further ligations were carried out where two separated linear DNAs were ligated together to form a plasmid

- PCR product was setup to self-ligate as follows: 8.5 µL of purified DNA, 1 µL of ligase buffer, 0.5 µL of T4 ligase enzyme. All ligations were incubated overnight at room temperature.

- Gel Purified fragments (2.5kb/7.8kb) were ligated overnight as follows: 1 µL of pSVP2741Aa 7.8kb fragment (ORF3), 3 µL of insert DNA 2.5kb fragment (ORF2), 0.5 µL ligase enzyme, 1 µL of ligase buffer, and 4.5 µL of distilled water were mixed and incubated overnight at room temperature.

3.2.5 DNA digestion

- Removal of methylated DNA

  DpnI enzyme was used to remove methylated DNA prior to transformation with bacterial cells. A total of 1 µL of DpnI was added to 45 µL of PCR product. This mixture was incubated for 1 h at 37C.

- Double digest and Isolation of ORF2 (2.5kb fragment) and ORF3 (7.8kb fragment)

  0.5 µL of BamH enzyme, 0.5 µL of XhoI enzyme, 2 µL of DNA, 1 µL of buffer2, and 6µL of distilled water were mixed and incubated for 1 h in a 37°C water bath.

The mixture was then loaded and run on an agarose gel, where desired bands were visualised under UV light and excised.

- HaeIII enzyme was applied to digest plasmids and confirm whether DNAs were desired constructs. Digestion tube contained: 1 µL of DNA, 7.5 µL of dH_2O, 1 µL of Buffer C, 0.5 µl of HaeIII enzyme. Tube was incubated for a minimum of 30 min at 37°C. Experimental restriction banding patterns were compared with the pattern predicted by the NEB web cutter.

3.2.6 DNA desalting

The salt was removed from DNA samples in preparation for Bt transformations. A 0.1M of glucose in 1% agarose gel was made as follows: 0.54g of glucose, 0.30g of agarose, and 30mL of distilled water were mixed and heated. The solution was aliquoted into eppendorf tubes and allowed to solidify. To desalt, 20-30µL of DNA was added to set agarose and allowed to sit for 30mins.

3.2.7 Bacterial transformation

A 100 mL broth solution of bacterial strain was incubated until OD600 0.4. sample centrifuged in SLA1500 at 10 rpm for 10 min. The pellet was washed several times in cold dH2O and finally aliquoted into 50 µL volumes. A 1µL of DNA was added the mixture and transferred to electroporation cuvettes. Gene pulser was set to 1.8Kv, 200ohms, 25 µF. The mixture was plated onto ampicillin (100 µg/mL) treated LB agar plates and incubated overnight at 37°C.

3.2.8 Bacterial Plasmid Miniprep using QIAprep Spin Miniprep Kit Protocol

An eppendorf tube with bacteria in 250 μL of buffer P1 was prepared. A total of 250 μL of buffer P2 was added, followed by, 350 μL of neutralising buffer N3. The mixture was centrifuged, and supernatant transferred to column. This was spun and flow through was discarded. A total of 500 μL of PB buffer was added to bound DNA and spun followed by and 750 μL of PE buffer. The DNA was eluted with 50 μL of EB buffer.

3.2.9 Rapid Size Screen (RSS)

RSS solution was prepared as detailed in section 3.1.2. Individual colonies were pick suspended in 25 μL of RSS solution. The samples were spun, and supernatant was loaded onto 1 % agarose gel. Colonies with potential mutants were selected for further analysis.

3.2.10 Protein harvesting

After 3days of incubation at 30 $^\circ$C, Bt cells were analysed under a light microscope (Leica DMLS). The cells were harvested once spores and crystals were observed. Cells were scraped from agar plates and suspended in 30 mL of cold distilled water. The cells were lysed by repeated sonication at 150 Watt (four cycles of 1 min) on ice. The pellet was collected by centrifugation at 12,000 x g for 10 min at 4 °C and re-suspended in 1 mL of distilled water.

3.2.11 Protein analysis using SDS-PAGE

Protein gels for this project were set up as follows unless otherwise indicated. RGB, SGB and 2x SDS loading buffers were made according to section 3.1.2. Reducing SDS/ME was made up of 95 µL of x2 SDS buffer and 5 µL of 2.5 % reducing β-Mercaptoethanol (ME). A total of 10 µL of protein sample was added to 10 µL of reducing SDS/ME and boiled for 3- 4min. The samples were spun and loaded onto SDS PAGE gel at 120 v for 30 min. Gel stain was made according to section 3.1.2. The gel was stained for 20 min and later treated with gel destain. Molecular masses of proteins were estimated by comparison with molecular standard protein marker (new England Biolabs).

3.2.12 Solubilisation and activation of Protein

Crude protein was spun, and pellet resuspended in carbonate buffer. Protein crystals were solubilised in 5x volume of carbonate buffers made according to section 3.1.2. One µL of 0.1 M of DTT was added to 20 µL of sample giving a final concentration of 5 mM of DTT. The mixture was incubated for an hour in 37 °C water bath. The samples were, spun and supernatant retained. Solubilised sample were analysed on protein gels. A total of 10 mg/mL of trypsin solution in carbonate was used for activation at the ratio of 1:10 (v/v) enzyme to supernatant and incubated in a 37 °C water bath for 1 h. Complete mini EDTA-free protease inhibitor was later added to stop further proteolysis.

3.2.13 Dialysis of Protein

Dialysis is designed to eliminate small molecular weight substances that might interfere with a subsequent step in the experimental procedure. It also used when large molecules such as proteins require different buffer environment for an experimental step (Glick and Pasternak, 1998). Here it was carried out in preparation for protein purification with AKTA –FPLC. The protein was dialysed against 1L volume of different Tris concentrations of pH 8. The protein was first dialysed against 1 L of 10mM Tris was made by dissolving 1.2 g of Tris-HCL solute in 1 L of water and pH adjusted with hydrochloric acid (HCL). The solubilised protein was loaded into a 350 µL volume capacity microanalysis button and covered with dialysis tubing which retains proteins with a molecular weight of 12,000KDa and above.

Large scale dialysis was carried out with 50mM Tris made by dissolving 6.0 g of Tris –HCL solute in 1 L of water. The solubilised protein was loaded into a 5mL capacity dialysis tube which retains proteins with a molecular weight of 12,000KDa and above. All samples were dialysed overnight in cold room gently stirred by magnetic stirrers.

3.2.14 Protein Purification

- Ion Exchange chromatography

    Solubilised dialysed protein toxin was purified using AKTA FPLC according to manufacture protocol. The following cold solutions were prepared and filtered overnight in cold room. One litre solution of 50mM Tris pH 8 was made by 6.0 g

of Tris-HCL solute in 1 L of water and split into two 500 mL flaks. A 500 mL of 1M sodium chloride in Tris was made by dissolving 30 g of NaCl in 500 mL of 50 mM Tris. A 500 mL of water and a 500 mL of 50 mM Tris were also prepared. A Millex GP micro-pore filter of 0.22 µM/ml was used to filter protein sample as it was injected. This was transferred to an anion exchange column Resource Q (amersham pharmacia biotech, UK). The toxin was eluted with an increasing linear gradient of NaCl in Tris at a flow rate of 1 mL/min. Fractions were collected and analysed on protein gels.

- Gel filtration

One mL of the solubilised trypsin treated protein was loaded to a Sephacryl S-200 High Resolution (Amersham) gel filtration (15 mL) column. The activated recombinant protein eluted with 50 mM sodium carbonate (pH 10.5). Fractions collected and quantified using BioRad kit. Fractions were also analysed on 7.5 % SDS-PAGE gels.

3.2.15 Protein concentration

Protein concentration was determined by the Bradford method (Bradford, 1976) using a Bio-Rad Protein Assay Kit (Bio-Rad) with BSA as the standard. Coomassie Brilliant Blue G-250 was added to a protein solution, which in acidic conditions binds to basic amino acids and can be measured at 595 nm with a spectrophotometer. The mixture was

incubated for 5 - 10 min at room temperature before measurement. Concentration of the unknown sample was determined by comparing its absorbance value against a plotted BSA standard curve. The standard curve showed near linear response (R^2=0.9975) over 0 - 1 mg/mL BSA concentration range.

SDS PAGE gel was also used to estimate protein concentration. Protein samples were loaded onto gel and the band intensities were compared against each other. The samples were diluted or concentrated by Vivaspin 20 centrifugal concentrator MWCO 100 kDa until bands appeared to be the same intensity. Further analysis of band density was carried out by BIO Rad software programme which measure band density and compares it to a control band density. It also estimates the molecular weigh of proteins against a protein marker in the same gel. Three such gels were analysed in this manner and average results are listed in results chapter 6.

3.2.16 Microscopy

Cell viability observation and post-treatment monitoring of swelling were carried under Nikon Eclipse TS100 inverted microscope.

3.2.17 Cell culture conditions

The cells were maintained under the conditions recommended by the supplier. For adherent cells, culture medium DMEM with high glucose (4.5 g/L) was supplemented

with 10 % FCS, and with 1 % PSG (100 U/mL penicillin, 100 µg/mL streptomycin, 292 µg/mL L-glutamine). Suspension cells were cultured in RPMI 1640 supplemented with 10 % FCS and 1 % PSG. For the culture of BL31 cells, RPMI 1640 was additionally supplemented with sodium pyruvate (100 nM) and 1-thioglycerol (100 nM).

The cells were cultured in sterile polystyrene 75 cm^2 flasks (Nunc) in standard cell culture conditions (37 °C and 5 % CO_2 humidified air). Medium was changed twice a week. When adherent cells reached confluency (between 70-80 %) they were washed with DPBS and detached by trypsinization (trypsin/EDTA containing 0.05 % trypsin and 0.53 mM EDTA) at 37 °C for 5 - 10 min. This was followed by centrifugation (100 - 200 x g for 5 min). Adherent cells were dispersed by banging and re-suspension in fresh medium. Cells were counted and split depending on the desired seeding ratio.

For electrophysiology experiments HepG2 cells were cultured in DMEM and low glucose (1 g/L). HeLa cells were cultured in MEM. Both mediums were supplemented with 10 % FCS and 1 % PS. Media were buffered with 25 mM Hepes and cultured in the absence of CO_2 in non-coated 60 mm plastic dishes.

3.2.18 Cell assays

Fluorescent cell assays were carried out in 96-well plates. Each well was loaded with 90 µl of cell suspension at a density of 5,000 or 22,500 cells per well (equivalent to 5.5 x 104 or 25 x 104 cells/mL) in complete DMEM (DMEM supplemented with FCS and PSG)

and incubated overnight under standard cell culture conditions. After incubation, 10 µl of toxin was added to each well. control wells contained 90µl of cell suspension and 10 µl of the appropriate buffer were set up. The following control buffers were used: for gel filtered toxin - 25 mM tris-HCL, 150 mM NaCl (pH 7.4), for ÄKTA purified toxin - PBS (7.4 pH). For background fluorescent control 100µl of cell culture medium was loaded into well. An additional control containing 90 µl of medium and 10 µl of toxin was also loaded into well. Generally, each treatment was tested in triplicate (three technical replicates). The viability of the cells was measured using CellTiter-Blue, usually after 24 h of treatment. For CellTiter-Blue assays cell density and incubation time with the reagent were optimised for each cell line. Twenty µl of the reagent was added at the end of toxin exposure period and the reading was taken after additional 2 h (end-point method). Fluorescence was measured with a green filter with excitation wavelength at 525 nm and emission wavelength range of 580-640 nm.

CellTox-Green assays were used for evaluation of cytotoxicity by measuring changes in membrane permeability. For CellTox-Green assays 'Express, No-Step Addition at Seeding Method' was selected. In brief: to 5 mL of cell suspension 10 µl of CellTox Green dye was added and then 90 µl of the mixture was dispensed into each well and allowed to adhere overnight. The next day each well was loaded with 10 µl of toxin. Fluorescence was measured before toxin addition and at regular 30 min time-points post-toxin addition using a blue filter with excitation wavelength at 490 nm and emission wavelength range of 510-570 nm. CellTox-Green values before toxin addition were subtracted from the

readings after treatment in that same well. Background fluorescent signals were subtracted from each well.

3.2.19 Statistical analysis

Results were expressed as means ± SEM for each experimental triplicate. To establish Cry41Aa and Y514A LC_{50} in HepG2 cells statistical analysis was carried out using SPSS software version 22.0, Probit Regression analysis (IBM, 2013). Dose response curves were claculated using nonlinear regression analysis and the equation 'Dose response Stimulation, log (agonist)vs. response-variable slope' in GraphPad prism version 7 for windows, GraphPad software, La Jolla California, USA, www.graphpad.com.

3.2.20 Planar lipid bilayer

A lipid mixture was created that contained phosphatidylethanolamine (PE), phosphatidylcholine (PC) and cholesterol (Avanti Polar Lipids) in a ratio of 7:2:1 (w/w). first, lipids were dried with N_2 to evaporate chloroform and then dissolved in n-decane, at a final concentration of 20 µg/mL. 1 mL of KCl buffer (150 mM KCl, 1 mM CaCl2, 10 mM Hepes, pH 7.5) was added to each of the artificially made disposable chambers. A 250 µm diameter hole present at the junction of both chambers was painted with the lipid mixture using a blunt end glass pipette. However, only the central area of ˜ 100 µm in diameter created a functional bilayer. Chambers were connected to electrodes through conducting agar bridges and located inside a Faraday cage to decrease electrical interference. Voltage was applied to the cis chamber (trans chamber was grounded).

Membrane activity was recorded for 30 min before toxin(s) was added. Capacitance of the membrane was between 180 - 200 pF and the painted membranes remained stable for a couple of hours. Cry41Aa and F509A were dosed at a concentration between 4-8 µg/mL, the current was recorded at different applied voltages. At the end of each of the three experiments, concentration of KCl in the cis chamber was increased to 450 mM to test if the pore was more cation or anion selective.

All experiments were performed at room temperature. Voltage was applied using Axopatch – 1D (Molecular Devices). Detected current was recorded and amplified by the same instrument and converted into a digital signal with Axon Digidata 1440A (Molecular Devices). Currents were filtered at 5 kHz and digitized at 50 kHz. Recordings were analysed using pCLAMP 10.5 (Molecular Devices). Conductance was measured by recording 20 to 25 current jumps for each voltage and averaged. Currents were plotted versus applied voltages (I/V curves) and the data points were fitted by linear regression. Conductance (G) was read from the slope of each I/V curve regression line. Probability of channel opening was evaluated from recordings by digital analysis of the distribution of all current values and plotted for each voltage as a histogram using pCLAMP 10.5 software.

3.2.21 Patch clamping

A day before the experiments, cells were seeded at a low density on 24 mm circular glass coverslips inside 35 mm dishes, prior to patching, cells were washed three times with buffer and mounted inside a coverslip holder that was fitted to the stage of an inverted microscope (IMT-2, Olympus). 1 mL of either NaCl (140 mM NaCl, 5 mM KCl, 1.1 mM $MgCl_2$, 1.1 mM $CaCl_2$, 10 mM Hepes, pH 7.4) or KCl buffer (140 mM KCl, 1.1 mM $MgCl_2$, 0.1 mM EGTA, 10 mM Hepes, pH 7.4) was added.

Pipettes were prepared using a Narishige glass capillary puller in a two-step heating process that required two types of pipettes. The borosilicate LG16 (Warner Instruments) for some of the single channel recording and soda lime 200 (Kimble Chase) for single channel and whole cell recordings. The heater was set to 84/77 for borosilicate and 84/74 for soda lime pipettes.

The following steps were used in patch clamp experiments: cell attached, whole cell, inside out. Pipettes were mounted inside a holder attached to an automatic micromanipulator. Ag/AgCl pellets created an electrical connection between the pipette and the amplifier. The bath solution was grounded. In some experiments, the pipette was first filled with buffer, followed by toxin(s) addition on top to allow slow diffusion towards the tip of the pipette. In other experiments, the toxin(s) was added straight to the bath. The filled pipettes had the resistance of about 4 MΩ and the resistance of the

seal was in the range of 2 - 10 GΩ. A voltage of -20 or -40 mV was applied to facilitate the transition from the attached cell into the whole cell mode.

Currents were launched as a set of seventeen 1 s depolarizing potential steps from −20 to 140 mV (10 mV increments). Current values between 0.95 – 1 s were analysed at each voltage, counting from the launch of each recording. Values were corrected for the averaged baseline current measured just prior to the applied potential. Current – voltage curves in the whole cell experiments were generated from each data set for each time point and conductance was calculated from the slope of regression line of each I/V curve.

In the single channel experiments, with excised patches conductance was calculated as described for PLB experiments using the slopes of the I/V curves. The following equation was used for cell attached mode: G=I/(V-Vr) to correct for the intracellular voltage, were Vr is a reversal potential. After plotting data points, Vr was calculated from the intersection of the voltage axis with the linear regression at zero current.

All experiments were performed at room temperature. Voltage was applied using Axopatch – 1D (Molecular Devices). Detected current was amplified and converted into a digital signal with Axon Digidata 1550 (Molecular Devices). Currents were filtered at

10 kHz and digitized at 50 kHz. Recordings were analysed using pCLAMP 10.6 (Molecular Devices).

3.2.22. Western blots

Cells were seeded in 6, 12, 24-well plates. the cells were lysed with RIPA, NP-40 (1 mL of lysis buffer per 2.8 x 107 cells) or 1.5% NOG (1 mL of lysis buffer per 2 x 107 cells). Each lysis buffer contained protease inhibitors: Complete protease inhibitor mixture, 1 mM EGTA, 1 mM EDTA and phosphatase inhibitors: 2 mM sodium orthovanadate and 1 µM microcystin. HepG2 cells were seeded at the density of 25 x 104 cells/mL in 6 well-plates (75 x 104 cells/dish). Sodium arsenite was added once cells reached >70% confluency. They were then incubated for 15-30 min with buffer, or toxins followed by 2 washes with DPBS. Complete cells were scraped, spun at low speed (200 x g) and supernatant was discarded. Pellet containing complete cells was re-suspended in 100 µl of NP-40 with protease and phosphatase inhibitors (1 mL of lysis buffer per 7.5 x 106 cells). Samples were vortexed for 10 s, left on ice shaking for 20 min and spun at 16,873 x g for 15-30 min at 4°C. Supernatant was collected and diluted protein concentration was assayed by Bradford.

For a p-p38 western blot 5-30µg of protein were loaded per well. The proteins were resolved on SDS-PAGE gels. The gel was drenched in dry blot buffer and proteins were transferred to a nitrocellulose membrane (HyBond ECL or Bio-Rad, pore size 0.45 µm)

via a Bio-Rad Trans-Blot Semi-Dry Transfer Cell system. The membrane was washed with PBS and blocked with PBS containing 0.02% Tween-20 (PBS-T) and 3% BSA for 1-3 h at room temperature. Further membrane washes with PBS-T.

Probing was done overnight at 4°C by incubating the membranes with rocking in a fresh blocking solution containing 3% BSA and primary antibody diluted in PBS-T (1:1000 dilution for antibodies against total and phosphorylated ERK and p38; 1:50000 dilution for anti-CD59 antibody). In pCREB and N-cadherin western blots both blocking and incubation with primary antibody (1 μg/mL final) was done using 5% non-fat dry milk instead of BSA. Post overnight incubation, the membranes were washed 3 x 10 min with PBS-T, followed by 1 h incubation with an appropriate horseradish peroxidase-conjugated secondary antibody diluted 1:2000 in PBS-T containing 5% non-fat dried skimmed milk. lastly, the membranes were washed again (3 x 10 min, PBS-T) and saturated in chemiluminescent detection solution and exposed to X-ray film (FUJI medical X-ray film).

The following antibodies were acquired: Rabbit monoclonal antibody against phospho-p38 (Thr180+Tyr182) and rabbit polyclonal antibodies against phospho-ERK, total p38 and total ERK were from Cell Signalling Technology (9215S, 9101S, 9212 and 9102 respectively). Second vial of rabbit monoclonal antibody against phospho-p38 (Thr180+Tyr182) was obtained from Thermo Fisher Scientific (MA5 15182). The

following primary antibodies were purchased from Abcam: rabbit monoclonal against

CD59 (ab126777).

Figure 13 Methods schematic diagram
Schematic diagram of the steps, and experimental procedures taken to create Cry41Aa constructs

Design Mutagenesis

Decide on mutagenesis, such as deletion, substitution, or creating a hybrid.
Mutagenesis took place in domain II loops of Cry41Aa, located in ORF2 of pBS41Aa
Design primers for mutagenesis.
Prepare template and PCR mix; optimize annealing temperature for PCR.
Run PCR products on 1% agarose gel, excise, and purify product. Allow linear DNA to self-ligate overnight at room temperature.

Transformation with *E. coli*

Ligation mixuture are introduced into *E. coli* and plated onto 100µg/ml ampicillin plates.
Colonies were picked and incubated in LB with ampicillin.
Plasmids were extracted using mini prep extraction kit

Confirmation of mutagenesis in *E. coli* expression

Plasmids undergo a Hae III digest and fragments are visualised on 1%agarose gel
NEB cutter software provides Hae III digestion profile for each mutant plasmid. These were compared against gels of plasmid Hae III digests. Mutant plasmids were further confirmed by sequencing.

Creating *Bt* expression vector

pBS41Aa lacks ORF3 of Cry41Aa required for packaging and crystal formation. Mutant OFR2 were isolated and taken into a *Bt* expression vector pSVP2741Aa.
Two double digests were carried out (i) BamHI and XhoI double digest to isolate mutant ORF2 fragment (2.5kb) from Blue Script.
(ii) a BamHI and XhoI double digest to isolate ORF3 from pSVP27*41Aa*. This is 7.8kb fragment. All fragments were visualised, excised and purified from agarose gel.
The liner DNAs were ligated overnight to form construct.

Transformation with *E. coli*

Ligation mixture were introduced into *E. coli* and plated onto 100 µg/ml ampicillin and 5 µg/ml chloramphenicol agar plates.
Colonies harbouring potential mutants are picked and incubated in LB with antibiotics.
Plasmids were extracted by mini prep kit and subjected to further analysis

Confirmation of mutagenesis in *Bt* expression vector

Hae III digests profiles on potential construct revealed potential candidates on 1.5% agarose gel.
Sequencing analysis confirmed desired construct.

Transformation with *E. coli*

Constructs in *Bt* shuttle vector was introduced into *E. coli* GM2163 to remove methylated DNA. Transformation was plated onto 100 µg/ml ampicillin agar plates. Colonies were picked and incubated in LB with antibiotics. Mini preps were carried out to isolate plasmids. Hae III digest were carried out on recovered plasmids and run against control DNAs on 1.5% agarose gel.
Constructs were confirmed by Hae III digests.
Construct were ready for transformation with *Bt*4D7.

Transformation with *Bt*4D7

The DNA is desalted to improve rate of successful *Bt* transformation.
Mutant constructs in *Bt* shuttle vector were introduced with Bt4D7 and plated out in 5 µg/ml chloramphenicol agar plates for overnight incubation at 30˚C. Colonies were picked and incubated in LB with 5 µg/ml chloramphenicol. The plasmids were extracted with mini prep kit for further analysis. The remaining LB with colony was plated out in 5 µg/ml chloramphenicol agar plates for 3 days at 30˚C.

Confirmation of *Bt* crystals and mutant DNA

Extracted plasmids were first visualised in agarose gel to confirm that DNA was obtained from *Bt* cells.
The plasmids were introduced into *E. coli and* plated on agar prepared with 100 µg/ml ampicillin and 5 µg/ml chloramphenicol and incubated overnight at 37 ˚C. Cell lawn was scraped, and DNA was extracted with mini pre-kit. Hae III digests confirmed that DNA which was introduced with *Bt*4D7 is the same one that was extracted from *Bt*4D7. Hae III digest were not carried out directly in *Bt* recovered plasmids as the presence of native plasmids digest profile would make HaeII digest profile of construct difficult to interpret.
Light microscope observations confirmed the presence of crystals and spore from *Bt* after 3days of incubation.
Bt cells were harvested and crystals were characterised

4.0 Bioinformatic analysis of Cry41Aa

4.1 Introduction

To study the function and structure of Cry proteins bioinformatic tools were used, which resulted in a growing nomenclature system that hosts over 300 Cry proteins (Crickmore *et al.*, 2018).

Key characteristics have been revealed, for instance many Cry proteins have five conserved sequence blocks (Höfte and Whiteley, 1989). In addition, molecular modelling has shown that Cry proteins with this feature tends to have a similar 3domain fold (Adang, *et al.,* 2017; de Maagd *et al.*, 2003; Pardo-López,2013; Soberón and Bravo, 2013).

Cry proteins tend to have high occurrence of homologous genes, however, the active form of the Cry protein tends to be active against a limited number of targets (Schnepf *et al.*, 1998; Crickmore *et al* ., 2013; de Maagd*et al.*, 2003; Bravo, 1997). Research studies have taken advantage of the specificity of Cry proteins and their ability to target certain insect orders to commercially explore their application as bio-pesticides. The ability of Cry toxin to recognise and to bind to different types of cells is a determinant of its specificity.

Studies on the adaptive evolution of Cry proteins have revealed potential positive selection for certain residues and positions that increase the diversity of specificity of Cry toxins. It was suggested that this increase on the type of target cell allowed Cry toxins to be active against a wider range of hosts where spores can germinate, ensuring their continued existence in the environment and their ability to overcome resistant adaptations of hosts (Wu *et al.* , 2007; de Maagd*et al.*, 2003).

A study by Mizuki *et al.* (1999) attempted to address the high rate of occurrence of Cry proteins without a known insect target cell. As a consequence, this class of Cry protein were named parasporins and allocated Cry protein names. Parasporins are non-haemolytic but active against certain cancer cell lines (Kitada *et al.*, 2005; Mizuki *et al.*, 1999; Ohba, Mizuki and Uemori, 2009; Mizuki *et al.*, 2000; Akiba *et al.*, 2009; Yamashita *et al*. 2005).

Among the parasporins, parasporin 3 and 5 are the least studied (Nagamatsu *et al.*, 2010; Yamashita *et al.*, 2005). Yamashita *et al.* (2005) first reported Cry41Aa also known as parasporin 3 as a 3-domain Cry protein of *Bt* that has cytocidal activity on myeloid leukaemia (HL60) and liver cancer (HepG2) cell lines. Molecular modelling suggests it has a secondary structure that is closely associated with the typical 3-domain structure of many insecticidal Cry proteins.

Cry41Aa is expressed from two open reading frames ORF2 and ORF3. Sequence alignments of ORF2 against protein sequences of revealed Cry structures has shown that the protein sequence of Cry41Aa encompasses the five conserved characteristic of insect targeting Cry proteins. Furthermore, ORF3 was found to contain conserved block 6, 7 and 8 which are also present in some larger insecticidal Cry proteins (Yamashita, 2005; Kitada *et al.*, 2005; Yamashita *et al.*, 2005; Nagamatsu *et al.*, 2010). Despite these similarities, Cry41Aa has no known insect target and its ability to identify, interact and kill HL60 and HepG2 cell lines points to a novel specificity not seen before in typical 3-domain Cry toxins (Kitada *et al.*, 2005; Ohba *et al.*, 2009).

Receptor recognition and binding specificity in 3-domain Cry proteins are usually associated with domain II (de Maagd *et al.*, 2001; Crickmore *et al.*, 1998; de Maagd *et al.*, 2003). It has been noted that domain I which is associated with membrane insertion and pore formation is highly conserved, allowing the majority of genetic divergence to occur elsewhere (Boonserm *et al.*, 2005; Schnepf *et al.*, 1998; Bravo., 1997). Another study has shown that mutations carried out in domain I resulted in reduced or total loss of toxicity (Saraswathy and Kumar., 2004).

Domain II is where much divergence occurs as it is associated with receptor recognition and binding specificity, and therefore its diversity within Cry proteins allows for a diverse range of target cells (Wu *et al.*, 2007). This is particularly true of the exposed loops typical of domain II, where loops 1, 2, and 3 vary in length and amino acid sequence

(Schnepf *et al.*, 1998; Smedley and Ellar, 1996; Tuntitippawan *et al.*, 2005; Adang, Crickmore and Jurat-Fuentes, 2014). Cross resistance by target cells is a consequence of high sequence homology of the loops in domain II among three domain Cry proteins which can bind to the same binding sites in midgut brush border (Adang, Crickmore and Jurat-Fuentes, 2014). A number of studies have identified residue positions in loop regions of domain II and domain III that are positively selected for, and mutagenesis resulted in loss or reduced toxicity (Wu *et al.*, 2007; Herrero *et al.*, 2004).

Other studies have investigated Cry toxin specificity and its improved specificity to target cell. These studies were also geared at the introduction of new specificity by addition to pre-existing one which resulted in the creation of novel toxins with the consequence of a quicker killing response (Wu *et al.*, 2000; Abdullah *et al.*, 2003; Pardo-López *et al.*, 2009; Dean. H and Sylvis L, 2006).

Pardo *et al.* (2009) carried out mutagenesis of loop2 and 3 on Cry1Ab to improve its insecticidal activity. It was discovered that these loops have important interactions with cadherin like transmembrane proteins CADR 11 and 12 binding sites and that the loops are critical for receptor interactions. Site mutagenesis on domain II loops of Cry3Aa and Cry1Ab, were reported to have a 10-fold increase toxicity of Cry3Aa towards its target insect *Tenebrio molitor*, and a 34-fold increase toxicity towards *Lymantria dispar* by Cry1Ab (Wu *et al.*, 2000).

Studies on domain II loops of Cry4Ba have shown that deletions and substitutions have resulted in a novel toxicity against *Culex* mosquitoes (Abdullah *et al.*, 2003). It is also possible to eliminate native toxicity and introduce novel activity against another target insect as illustrated by complete removal of Cry1Aa's activity to caterpillar larvae (Dean and Sylvis., 2006) and to introduce toxicity to *Culex pipiens* by replacing Cry1Aa's native loop1 and 2 with loop regions of Cry4Ba (Tuntitippawan *et al.*, 2005; Abdullah *et al.*, 2003; Dean. H and Sylvis L, 2006)

Early studies on loop 3 of domain II have implicated its role in target specificity, receptor recognition, and binding. Hussain *et al.* (1996) proposed that loop 3 residues have hydrophobic interactions with receptor and mutations in this loop can affect the toxins ability to bind. Indeed, Fujii *et al.* (2012) carried out research on conserved amino acids of Cry1Aa. The 30 amino acid loop that covers all its binding loops were substituted with hydrophobic residues and resulted with an increase in binding affinity towards cadherin $BT\text{-}R_1$ receptor (Hussain *et al.*, 1996; Fujii *et al.*, 2012).

Pacheco *et al.* (2009) demonstrated that mutations in domain II loops affect binding to the cadherin (Bt -R_1) and aminopeptidase-N (APN) receptors in the lepidopteran *Manduca sexta*. These studies point at the implication of loop 3 of Cry1Ab in its mode of action. Mutations in loop 3 of Cry1Ab were found to severely affect the toxin's insecticidal activity in both monomeric and oligomeric states of the activated toxin. A "ping pong" binding mechanism to explain the toxin activity and its dependency on loop

3 of domain II was proposed (Pacheco *et al.*, 2009). Such studies have highlighted the important role of domain II loops in the specificity of 3-domain Cry proteins.

In order to gain an understanding of the specificity that Cry41Aa has towards HL60 and HepG2 cell lines it is vital to explore the secondary structure of domain II loops In this study comparisons were made between the predicted Cry41Aa secondary structure and that of resolved 3-domain Cry protein structures. Bioinformatic tools were employed to carry out the comparative analysis.

4.2 Identification of the putative loops of domain II in Cry41Aa using bioinformatic tools

Information gathered from protein sequence and structure is often used to explore protein function. Sequence homology between proteins can suggest similar functions of proteins particularly if they share putative homologous regions and similar structural arrangements. Bioinformatic software uses a wide range of data gathered from known protein sequences and structures to present a model summary of an unknown protein sequence.

Information on protein sequence, structure, and function is ample. Some bioinformatics tools collate data to trace evolutionary patterns of proteins in an attempt to improve the understanding of protein function. Naturally, proteins diverge; domain swapping,

matching or mixing is another way of obtaining novel function. Bioinformatic tools aid to make sense of the vast array of information and collates it in a form so that it may predict protein function.

Scientists have employed computer programmes to assist with the comparisons and analysis of patterns and evolutionary relationships between biological sequences. Clustal is one of the bioinformatic tools that is comprised of computer programs for multiple sequence alignment of sequences. Protein, DNA, or RNA sequences can be aligned against a query or unknown sequence and depending on the algorithm; a sequence homology is detected and a phylogenic analysis is carried out to assess the links between the query sequence and that of known sequences. Multiple alignments are used to find diagnostic patterns to characterise protein families and determine homology between new sequences and existing families of sequences; in order to predict the secondary and tertiary structures of new proteins. Its takes advantage of the fact that homologous sequences are evolutionarily related.

The programme builds a multiple alignment progressively by a series of pairwise alignments, following the branching order in a phylogenetic tree. There are many versions of Clustal, operating systems and tools. It has evolved and progressed from its early 1988 version by Des Higgins. This study has applied ClustalW, the third version of a series of Clustal software. This version has sequence weighting, position-specific gap penalties and the automatic choice of a suitable residue comparison matrix at each stage

in the multiple alignment as well as more sensitive programming algorithm (Thompson *et al.*, 1994; Chenna *et al.*, 2003).

However, the accuracy and reliability of this output information varies from one tool to another (Whisstock and Lesk, 2003). Bioinformatics tools have narrowed down regions involved in specificity, receptor recognition, and binding by analysing unresolved Cry protein sequences against resolved Cry protein structure in attempt to understand Cry toxin mode of action (Li *et al.*, 1991). In an attempt to understand, which regions of Cry41Aa are responsible for its activity against HepG2, bioinformatics tools were employed as a starting guide in an effort to locate sequence(s) responsible for its specificity towards HepG2.

Clustal alignment analysis was applied to narrow potential loop regions in protein sequence of Cry41Aa. The protein sequences of resolved 3-domain Cry structures listed in table 6 were used in bioinformatics analysis of Cry41Aa protein sequence to identify its domain II loops.

List of resolved domain II loops

Cry toxin name	Domain II loop sequences		
	Loop 1	Loop 2	Loop 3
Cry1Aa	^{311}RG^{312}	^{368}RIILGSGPNNQE^{379}	^{435}TMLSQAAGAVYTLRAPT^{451}
Cry1Ac	^{311}RG^{312}	^{369}RPFNIGINNQQ^{379}	^{438}SGFSNSSVS^{446}
Cry2Aa	^{322}NIGGLPG^{328}	^{378}LDSGTDREGVA^{388}	^{440}RPLHYNQIRNIESPGTPGGAR^{461}
Cry3Aa	^{356}PGYYNDS^{363}	^{420}PS^{421}	^{489}MQGSR^{493}
Cry3Bb	^{349}PGYFGKDS^{356}	^{412}WPNGK^{416}	^{484}MQDRRG^{489}
Cry4Aa	^{371}KAQTTPNNF^{379}	^{460}AGSGQITYD^{468}	^{508}SIPATYKTQ^{516}
Cry4Ba	^{330}TI^{331}	^{410}SNITPTPEG^{418}	^{455}YN^{456}

Table 6 List of resolved domain II loops
 Table lists loops 1, 2, and 3 of domain II of the protein sequences of resolved structures of Cry proteins used for alignment against Cry41Aa.

The following is a list of crystal resolved Cry toxins: Cry1Aa (Grochulski *et al*., 1995), Cry1Ac (Evdokimov *et al.*, 2014), Cry2Aa (Morse *et al.*, 2001), Cry3Aa (Li *et al.,* 1991), Cry3Bb (Galitsky *et al.,* 2001), Cry4Aa (Boonserm *et al.,* 2006), Cry4Ba (Boonserm *et al.,* 2005), and Cry8Ea (Guo *et al.,* 2009).   An alignment of Cry41Aa against these Cry structures was carried out to locate putative loop region of domain II associated with specificity and shown in figure 14 a and 14 b (Likitvivatanavong *et al.,* 2009; Adang *et al.*, 2014; Crickmore *et al.*, 1998). The beta-trefoil domain of Cry41Aa was deleted from

the sequence before alignment as it has earlier been demonstrated that its removal has no impact to susceptible cell line (Krishnan, 2013). The alignment has highlighted the similarities between Cry41Aa and 3- domain Cry toxins. The alignment has been significantly narrowed down regions of putative loops of domain II in Cry41Aa where it aligned against known putative domain II loop of resolved Cry sequences.

The alignment was further analysed by Box shade where amino acids that are the same or similar are highlighted, increasing the confidence of predications made about Cry41Aa. From this, similar or identical amino acids are likely to have similar structural roles, and thus provide evidence of analogous structural role of Cry41Aa sequences as those of the resolved Cry sequences. The alignment highlights the distinctive extra loop not found in other 3-domain Cry proteins as noted in Figure 14 a and 14 b. It also indicates the poor alignment of Cry41Aa loop2 against the resolved Cry sequences when compared to how well loop1 and loop 3 of Cry41Aa align with loop 1 and loop 2 of resolved Cry sequences. Further tools were employed to identify the domain II putative loops of Cry41Aa.

```
cry2Aa    1 TDHS------------------LYVAPVVGTVSS LLKKV SLIGKRILSE WGI F S
Cry4Aa    1 QNQQYGGDFETF-IDSG-ELSAYTIVVGTV TGFG TTP-L LA--LIGFGT IPVLF A
Cry4Ba    1 NNQQYGVNPAAI---NSSSVS AL---KVAGAILK VNPPA TV--LTVLSAVLPIL T
Cry41Aa   1 SGDPTV GEG----YSADVRDAVITSINIASYLLSVPFFPA VA--AGILGA LGLL T
Cry1Aa    1 -IPYNC SNPEVEVLGGERIE GYTPIDIS SLTQ LLSEFVPG--AGFVLG VDI  GI
Cry8Ea    1 GYDSEYSGSPEVLISERDAVK AISLVGTI GKLGVPL--V PI--VSLYST IDVL -
Cry3Aa    1 DNNTEA DSS----TTKDVIQRGISVVGDL GVVG PF--G AL--VSFYTNFLNT -
Cry3Bb    1 DSSTEV DNS----TVKDAVG GISVVGQI GVVGVPF--A AL--TSFYQSFLNT -

cry2Aa   42 GSTN---LMQDILRET QFL RLNTDTLARVN  I  QANIREFNQQVDNFLNPTQNP
Cry4Aa   56 QDQS--NT SD IT TKNI KKE ASTYISN NKI NRSF VIST HNHLKT ENN NPQ
Cry4Ba   53 NTPTPERV ND  TNTGN D TVTA V TD N KMTVVKDYLDQ TTKFNT KRE NNQ
Cry41Aa  55 NTQA---V E  NTV A    KLD   SK IS N K VLEL QDAADD NEN GDL
Cry1Aa   58 FGPS---Q D  PV I Q   R E F NQ ISR E  S LYQI AESFRE EAD TNP
Cry8Ea   56 GGKS---Q EI  E V A   K A   AK L  E G NYQL LTALEE QEN SST
Cry3Aa   52 SE-D---P K   E V A MD K AD  RNK L   Q Q NVED VSALSS QKN VSS
Cry3Bb   52 SDAD---P K   A V V  DKK P   KSK L   Q Q NFED VNALNS KKT LSL

cry2Aa   99 VP----LSITSSVNTMQQL L RL Q- --------QIQG QL  PLF    M  SFI
Cry4Aa  114 NT----QDV TQIQLVHYH Q VI ELVNSCPPNPSDCDY NI V SS      TV
Cry4Ba  113 SY---RTAVITQ NLTSAKLRETAVYFS--------NLVG L  PI  V FN L I
Cry41Aa 112 RN---KNRVLTE RNVNGH E SM S---------RNF  N  PV P     L
Cry1Aa  115 AL---REEM IQ NDMN ALTTAI L-L--------QN Q P SV V     SV
Cry8Ea  113 RVLR---DV NR EILD L TQYM S--------R TG   F SV      L
Cry3Aa  108 RNPHSQGRI EL SQAE H R SM S--------- ISG  F IT     T F
Cry3Bb  109 RSKRSQDRI EL SQAE H R SM S--------- SKF  F PT     T L

cry2Aa  146  VILNAD   IS-----AATL---RTYRDY RNY RD SNY INT QTAFRG NTRL--
Cry4Aa  170 NQ VK EAYLKNNRQFDYLEPLPTAIDY PV TKAIED NY TT KK NLIKTTP--
Cry4Ba  162  GLINAQ  SLA-----RSAG---DQL NTMVQY KE IA SIT  K DV RNK---
Cry41Aa 160  VK EG MS-----TDPGAERDDM RR RSR EI    NT Q  QQAKSLQA N
Cry1Aa  163  VSW QR  FD-----AATI---NSR ND TR IGN  YA R  T ERVW P--
Cry8Ea  161 K SI E  FS-----TTAI---NNY NRQMS IAQ S  Q RT DR K S---
Cry3Aa  159 K QIY E  YE-----KEDI---AEF KRQLK QE  K V DK R S---
Cry3Bb  160 K QV E  YS-----SEDV---AEF HRQLK QQ   N V NG R S---

cry2Aa  196 ------------------------------------HDMLP TY F N FEYVSIWSLFKYQ
Cry4Aa  228 ------------------------DSNLDGNIN NT TY TK TA VV  N VG
Cry4Ba  211 ------------------------SNGQ ITF DYK  IQ L AS P
Cry41Aa 215 VSDYSRYPWTQYNQSGGFSYREAKGEYRGT EN NL A  D IL  I Q T PG
Cry1Aa  212 ------------------------DSRD R  Q L T  V SN S
Cry8Ea  210 ------------------------NAKQ E R    S MT M M
Cry3Aa  208 ------------------------SYES NF RY   T LI  L V
Cry3Bb  209 ------------------------TYDA KF R    T LIV F I

cry2Aa  223 SLMVS-------SG----ANL ASGSGPQ-QT----------QS T-------AQNW F Y
Cry4Aa  265 K IG-------VQS    QVLNFEESPYKYY-------D QYQ DS--LT R
Cry4Ba  243 R ADKIDNTKLS T F   ALVESPS----------SKSIAAL AA--LT DV
Cry41Aa 275 L SRP-------V S  U DIRGTT-----WRSDANLNTIDA NRMVGS QLQ
Cry1Aa  244 R IR-------TVSQ   N VLEN-FDGSFR-----GMAQR  -Q--N Q M
Cry8Ea  242 T ME-------T AQ  V D IGAIGAQGSWYDSA--PS NTL ST--F GK
Cry3Aa  240 L KE-------V T   DVI D IVGVNNLRGY---G--TT SN  -N--Y K
Cry3Bb  241 L SKG-------V T  D F D IFSLNTLQEY---G--PT LS  -N--S K

cry2Aa  255 SLFQVNSN ILSGI           NIG LP GST-----TT SLNSARVNY-----SGGVS
Cry4Aa  309 TW DSLN E-------------KAQTTPN-----NF TS YNM-----FHYTLDNISQK
Cry4Ba  291 TW KRVD M-------------NTI QD----L-----RFLSANKIGFSYTN SAMQE
Cry41Aa 323 TW TEMK I-------------RNTG SITS-----TH DLMV-----G----LEKKI-
Cry1Aa  288 I NSITI-------------DVH R T-----FN S QITAS VGF-----GPEFA
Cry3Aa  285 Y HRIQ H-------------RFQP YYGNDS FN S NYVSTR SIG----NDIIT
Cry3Bb  286 Y QGIE H-------------RLQP YFGKDSSYN S NYVETR SIG----SKTIT
```

Figure 14a ClustalW alignment analysis of Cry41Aa

Clustal alignment analysis of Cry41Aa against the protein sequences of revealed crystal structures of Cry2Aa, Cry3Aa, Cry3Bb, Cry4Aa, Cry4Ba, and Cry8Ea. Box shade secondary structure analysis highlights identical amino acids in black and similar amino acids in grey. A unique sequence distinctive only to Cry41Aa known as the extra loop is highlighted in pink named the extra loop of Cry41Aa with the following sequence [261]NVSDYSRYPWTQYNQSGGFSYREAKGEYRGT[291]. Loop1 is boxed in red for the known Cry toxins, and highlighted in red for predicted loop1 of Cry41A with the following sequence [384]SITS[387]. Loop2 is boxed in yellow for known Cry toxins, and highlighted in yellow for predicted loop2 of Cry41Aa with the following sequence [416]QNTSYTRIDRP[426]. Loop 3 is boxed in blue for the known Cry proteins, and highlighted in turquoise for predicted loop 3 of Cry41Aa with following sequence [503]VRDNCPFAWPGYKQ[L517].
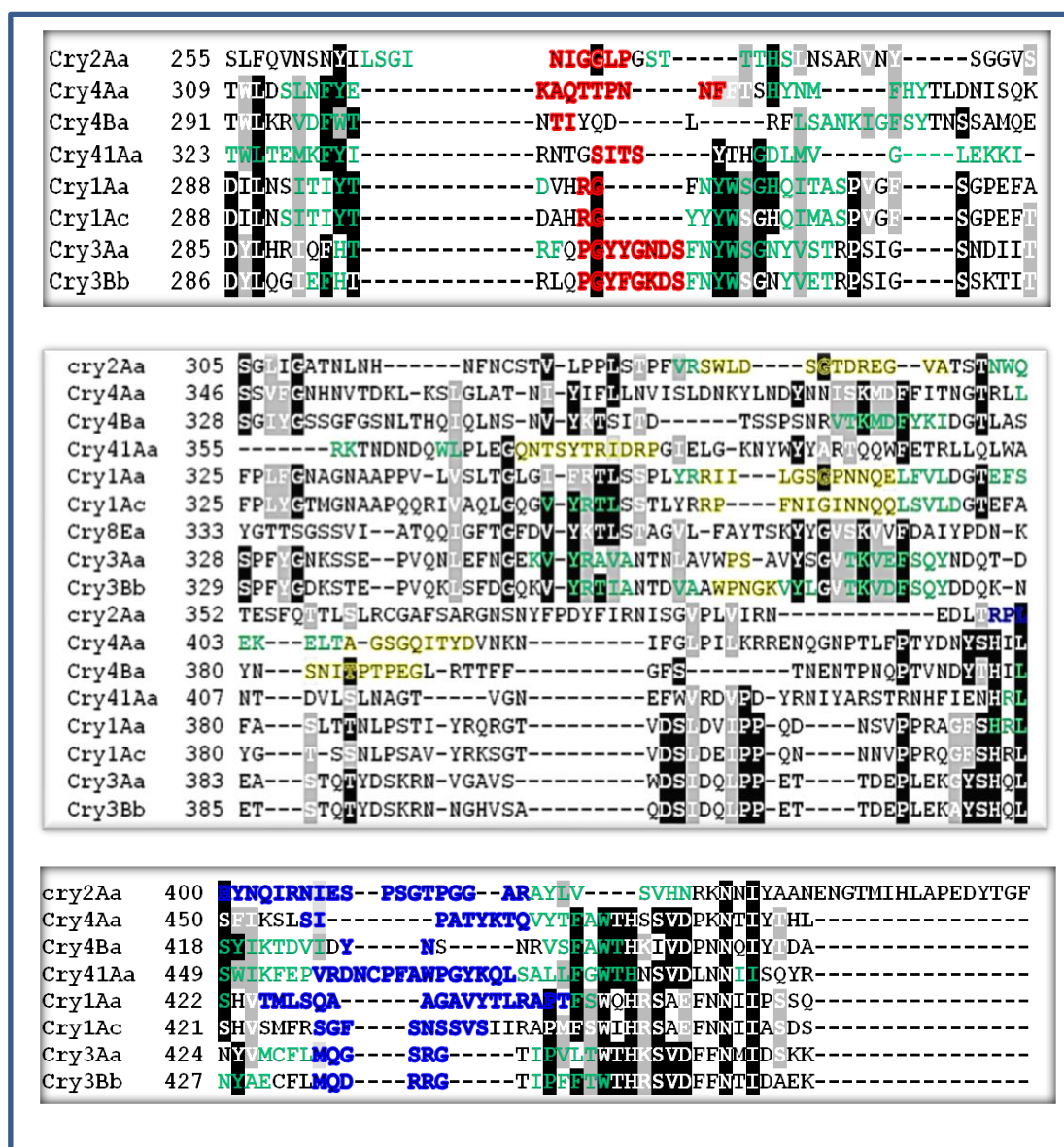
Figure 14b Analysis of Cry41Aa by ClustalW alignment and Box shade
ClustalW alignment analysis of Cry41Aa against the protein sequences of revealed crystal structures of Cry2Aa, Cry3Aa, Cry3Bb, Cry4Aa, Cry4Ba, and Cry8Ea. Box shade secondary structure analysis highlights identical amino acids in black and similar amino acids in grey. For all Cry proteins listed, Loop1 is highlighted in red, Loop2 is highlighted in yellow, loop 3 is highlighted in blue. β sheets are highlighted in green.
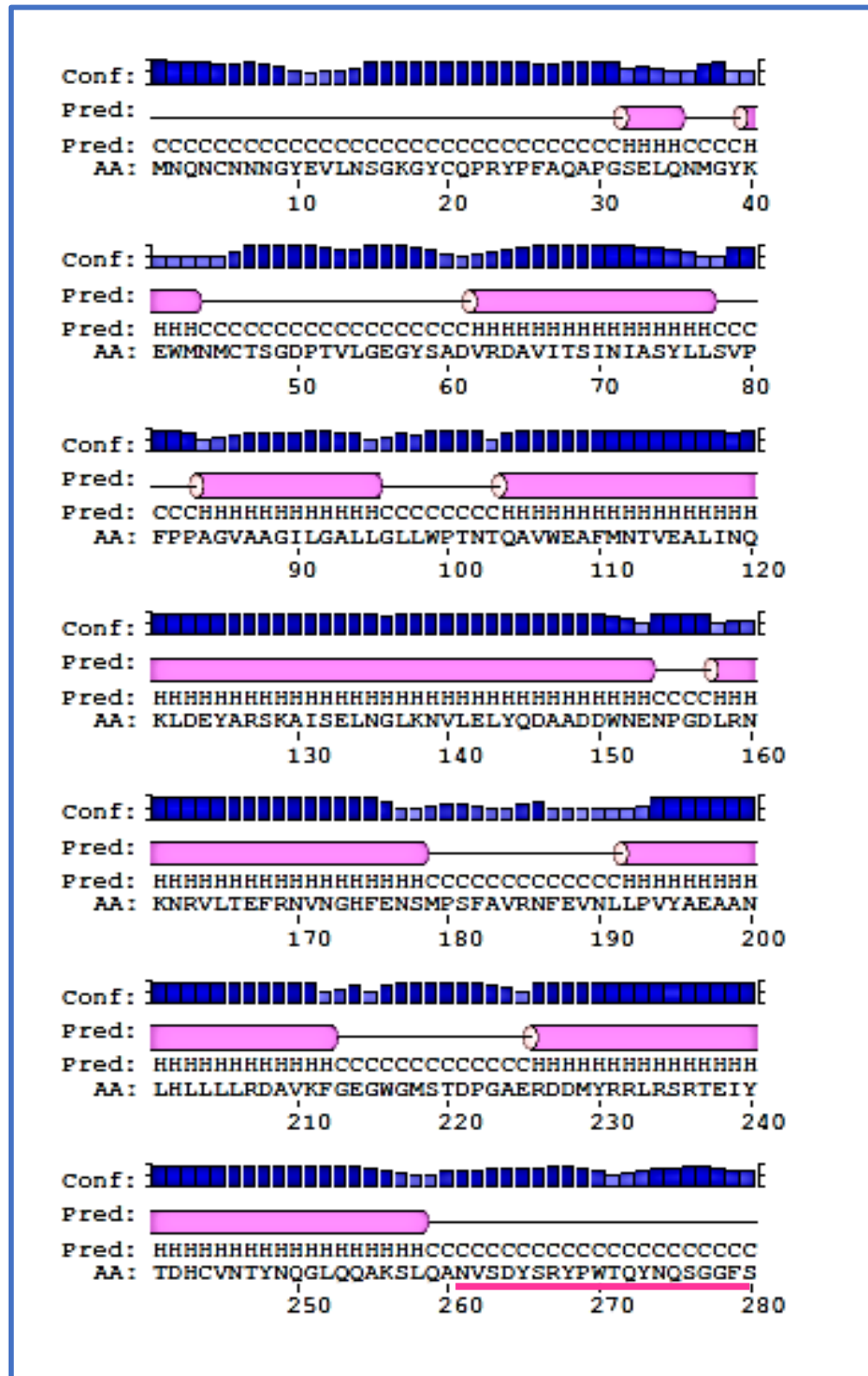
PSIRED web server predicted the secondary structure of Cry41Aa, where sequences are predicted as β- strands, α- helixes, or coils and given a level of confidence as the accuracy of its predication as shown in figure 15, and 16. The PSIRED protein structure

predication server incorporates three key programmes that predict secondary structures of globular and membrane bound proteins. In addition to secondary protein structure analysis, it incorporates the GenTHEADER program which recognises protein folds as well as the MEMSAT 2 program which predicts structure and topology of transmembrane proteins (McGuffin *et al.,* 1999).

The overall secondary protein is predicted to be highly helical in particular domain I. The analysis complements findings from the other bioinformatics tool and suggested that the three putative loops of domain II of Cry41Aa are found as coils structures. Loop 1 is predicated to sit in a coil between two β-sheets and is made up of four amino acids. Out of the three domain II loops, loop1 has the highest confidence score that is predicted loop structure is more likely. Loop 2 is the least confident predication of the three loop regions. It is made up of 11 amino acids found in a coil region made up of 31 amino acid. β-sheets here have the furthest proximity from loop region compared to the other two predicated loop regions. Loop 3 is predicated with a good level of confidence. It is made up of 15 amino acids flagged up by two β-sheets immediately either side of it. It is predicated to be in coil conformation made of a total of 17 amino acids that include the 15 amino acids that make up loop 3 detailed in figure 15. The extra loop is flagged by helical structures either side.

Figure 16 summarises the secondary structure of PSIPRED analysis. The three loops were mapped out onto the ClustalW alignment including the β-sheets that fall either side of

the three predicted putative of Cry41Aa as well these of the resolved 3-domain Cry

proteins.

```
Conf: }███████████████████████████████████████
Pred: ──────────────⊂████████████████⊃──────────
Pred: CCCCCCCCCCCCHHHHHHHHHHHHHHHHHHHHCCCCCCC
 AA: YREAKGEYRGTENWNLYNAFRRDMTILVLDIIAQFPTYDP
              290       300       310       320

Conf: }███████████████████████████████████████
Pred: ──────────────────⇒──────────────⊂█████⊃──
Pred: CCCCCCCCCCCCCCEEEECCCCCCCCCCCCCCCCCHHHHHHHC
 AA: GLYSRPVKSELTREVYTDIRGTTWRSDANLNTIDAIENRM
              330       340       350       360

Conf: }███████████████████████████████████████
Pred: ──────────────⇒════════⇒──────────────▭───
Pred: CCCCCCCCCCCCCCEEEEEEECCCCCCCCCCCCCCCCCEEE
 AA: VGSRQLQLFTWLTEMKFYIRNTGSITSYTHGDLMVGLEKK
              370       380       390       400

Conf: }███████████████████████████████████████
Pred: ⇒──────────────────────────────────────⇒
Pred: EECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCEEEEEEC
 AA: IRKTNDNDQWLPLEGQNTSYTRIDRPGIELGKNYWYYART
              410       420       430       440

Conf: }███████████████████████████████████████
Pred: ──────────────────⇒──────────────────────
Pred: CCCCCCCCCCCCCCCCEEEECCCCCCCCCCCCCCCCCCCCCC
 AA: QQWFETRLLQLWANTDVLSLNAGTVGNEFWVRDVPDYRNI
              450       460       470       480

Conf: }███████████████████████████████████████
Pred: ──────────────────⇒──────────────────▭───
Pred: CCCCCCCCCCCCCCCCEEEEEECCCCCCCCCCCCCCCCCCCEEE
 AA: YARSTRNHFIENHRLSWIKFEPVRDNCPFAWPGYKQLSAL
              490       500       510       520

Conf: }███████████████████████████████████████
Pred: ⇒──────────⇒──────⇒──────────────⇒
Pred: EEEEEECCCCCCCCEECCCCCEEEECCCCCCCCCCCEEEEEC
 AA: LFGWTHNSVDLNNIISQYRITQIPAVKAYWNRGAFSVIRG
              530       540       550       560
```

Figure 15 PSIRED webserver secondary structure image prediction of Cry41Aa based on amino acid sequence.
Loop1 is underlined in red, loop2 underlined in yellow, and loop 3 underlined in blue. The extra loop distinctive to Cry41Aa is underlined in pink (Jones, 1999; Buchan *et al.*, 2013).

Figure 16 PSIRED webserver secondary structure image prediction of Cry41Aa structure based on amino acid sequence.
Helical amino acids are highlighted in dark grey. β-sheets are highlighted in light grey.   Loop1 is boxed in red, loop2 in yellow, loop 3 in blue. The extra loop distinctive to Cry41Aa is boxed in pink

X-ray crystallography remains the main method for the determination of protein structure, however it is time consuming and it is successful only when specific conditions for growing crystals are optimised. Thus, the main method to predict protein structure takes advantage of computational homology modelling which can predict the secondary structure of a protein from its sequence (Shokry *et al.*, 2018).

Protein modelling servers SWISS-MODEL and the Protein Homology / analogY Recognition Engine V 2.0 or PHYRE2 were used to analyse the secondary structure of Cry41Aa. The PSIPRED analysis predicted the secondary structure based on the protein sequence of Cry41Aa highlighting α helixes, coils and β sheet strands in Cry41Aa with a score of confidence. Of the three loops, loop1 aligned better than any of the other loops predicted in the alignment.

Many homology-modelling methods involve template selection, target template alignment, model building, and evaluation of the model as steps in their modelling approach. The named processes are repeated and reassessed until a model structure is created. The SWISS-MODEL server is based on the rigid fragment assembly approach, where the framework of the unknown protein is calculated on the basis of the family of homologous proteins, or an appropriate subgroup (Schwede *et al.*, 2000; Schwede *et al.*, 2003). The program then aligns fragments of the known protein structures of high sequence homology with the unknown. This alignment provides a basis for building the model's the structure. The server implements four main steps: template superposition,

target-template alignment, model building and energy minimization. In brief, the server first makes a template selection by searching a template library database. Quality indicators such as empirical force field energy removes low matching templates. An alignment is then created. A local pair-wise alignment of the target sequence to the main template structure is calculated and a root mean square differences (RMSD) between pairs of homologous structures score can be obtained. A model of the unknown protein sequence is built using information from the averaged backbone atom positions of templates. Loops are built based on scoring schemes, that consider field energy, steric hindrance and interactions such as hydrogen bond formations. The model is complete once energy minimization is carried out. In this final step, parts of the model with conformational errors are detected. Empirical force fields are used by the SWISS-MODEL server to regularise the built model (Biasini *et al.*, 2014; Schwede *et al.*, 2000; Schwede *et al.,* 2003).

The Protein Homology / analogY Recognition Engine V 2.0 or PHYRE2 is also a protein homology modelling server. It applies homology modelling principles to build a model of an unknown protein sequence. It is based on the understanding that the structure of proteins is conserved in evolution and that a model of an unknown protein can be made from its sequence based on related known protein sequences or templates (Kelley *et al.*, 2009; Kelley *et al.*, 2015). The server applies four main stages to create a model. The server first scans the unknown protein sequence against a protein database where it gathers homologous sequences that have at least 20 % sequence identity with the query

sequence to build a multiple sequence alignment using the HHblits algorithm method. The resulting alignment is then assessed in PSIPRED a software that predicts the presence of α-helices, β-strands and coils with an accuracy of 75–80% (McGuffin *et al.,* 1999). It assesses both the predicted secondary structure and the alignment. HMMs output proceeds to create the backbone of the model. This is followed by loop modelling and finally side chain placement before a model is presented (Kelley *et al.,* 2009; Kelley *et al.*, 2015).

Two protein-modelling servers SWISS-MODEL and PHYRE2 were applied to predict the secondary and tertiary structures of Cry41Aa. Assisted by UCF chimera software to visualise the structures. SWISS-MODEL generates 3D models for proteins using a template library in order to put together complete structural models matching Cry1Aa as the closest homologue at the time of the analysis. (Biasini *et al.*, 2014).

The SWISS-MODEL server created a model of Cry41Aa using its 'first approach mode' (Biasini *et al.*, 2014). Here the sequence of Cry41Aa was screened against a number of suitable templates obtained from the ExPDB library. The automated modelling system then produced a model against a template that shares at least 25 % sequence identity with Cry41Aa. A number of insecticidal templates were identified against the sequence of Cry41Aa, however the template with the highest percentage of sequence identity was that of Cry1Aa with a 34 % sequence identity with Cry41Aa. it has a Global Model Quality Estimation (GMQE) score of 0.48. The GMQE is usually a number between 0 and 1 that

represents the accuracy of the alignment, template and built model. The higher the score the more reliable the alignment and thus the model built (Waterhouse *et al*., 2018). Figure 17 illustrates the SWISS-MODEL model of Cry41Aa. Similar to insecticidal Cry1Aa the SWISS-MODEL modelled Cry41Aa has a domain I is made up of α-helixes and domain II is a β-sheets sandwich both typical of 3 domain insecticidal Cry protein. Domain III is made up of β-sheets and a network of exposed loops, however Cry41Aa differs as it has a distinctive partial α- helical extra loop (shown in pink, figure 17) that appears to protrude away from the main structure. Here, loop1 is short and tucked under, loop2 and loop 3 appear to extend out from the main structure. The model and Cry1Aa template were superimposed and were calculated to have an RMSD value of 0.7.

PHYRE2 server applies template based homology modelling and fold recognition methods to predict secondary and tertiary structures to produce a PDB file for visualisation and analysis (Kelley *et al*., 2015) . The bioinformatic PHYRE2 tool has analysed the Cry41Aa sequence. It modelled 68% of the Cry41Aa sequence with 100% confidence. Furthermore, it suggested that the resolved crystal structure of Cry3Aa is the optimum structural template despite sharing 32% sequence identity with Cry41Aa sequence.

The PHYRE2 server has modelled Cry41Aa to be structurally similar to insecticidal 3-domin Cry protein such as Cry3Aa in figure 18. Domain I has the typical α-helices of

insecticidal Cry proteins. Domain II is made up of β-sheets sandwhich that appears to be in a looser confromation to that seen in Cry3Aa. Similarly, domain III of Cry41Aa is made up of β-sheets and a network of exposed loops. Here, the extra loop appears to be closer to the main structure and it is not helical but rather in a coil formation. Loop1 and loop 2 appear to be positionally similar to those of Cry3Aa. when the model was superimposed onto the Cry3Aa template is had a RMSD value of 0.6.
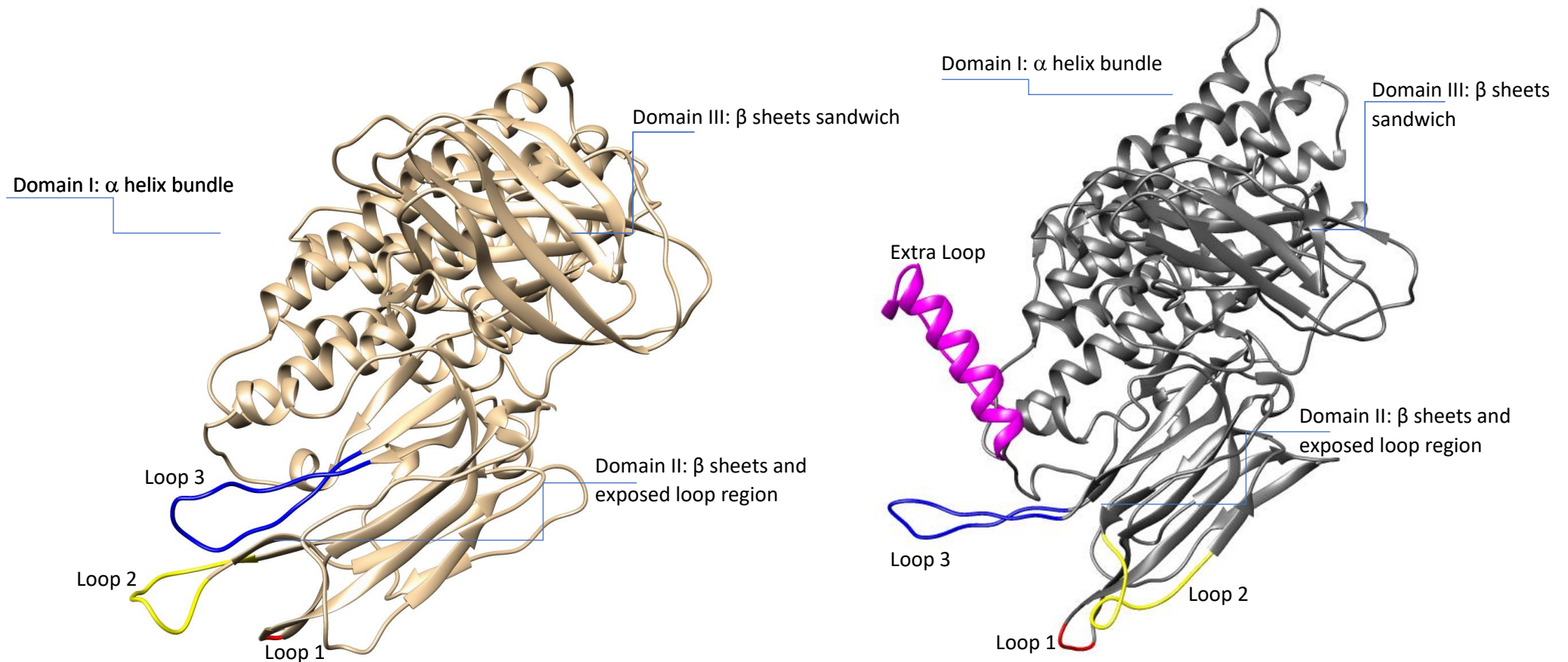
Figure 17 SWISS-MODEL visulaised Cry41Aa structure.
On the left is the 3D structure of crystallography resolved Cry1Aa shown in light grey. Loop1 is shown in red, loop2 in yellow, and loop 3 in blue. On thr right is the SWISS-MODEL generated predictaed strructure of Cry41Aa shown in dark grey. Loop1 is shown in red, loop2 in yellow, loop 3 in blue, and the extra loop exclusive to Cry41Aa in pink.
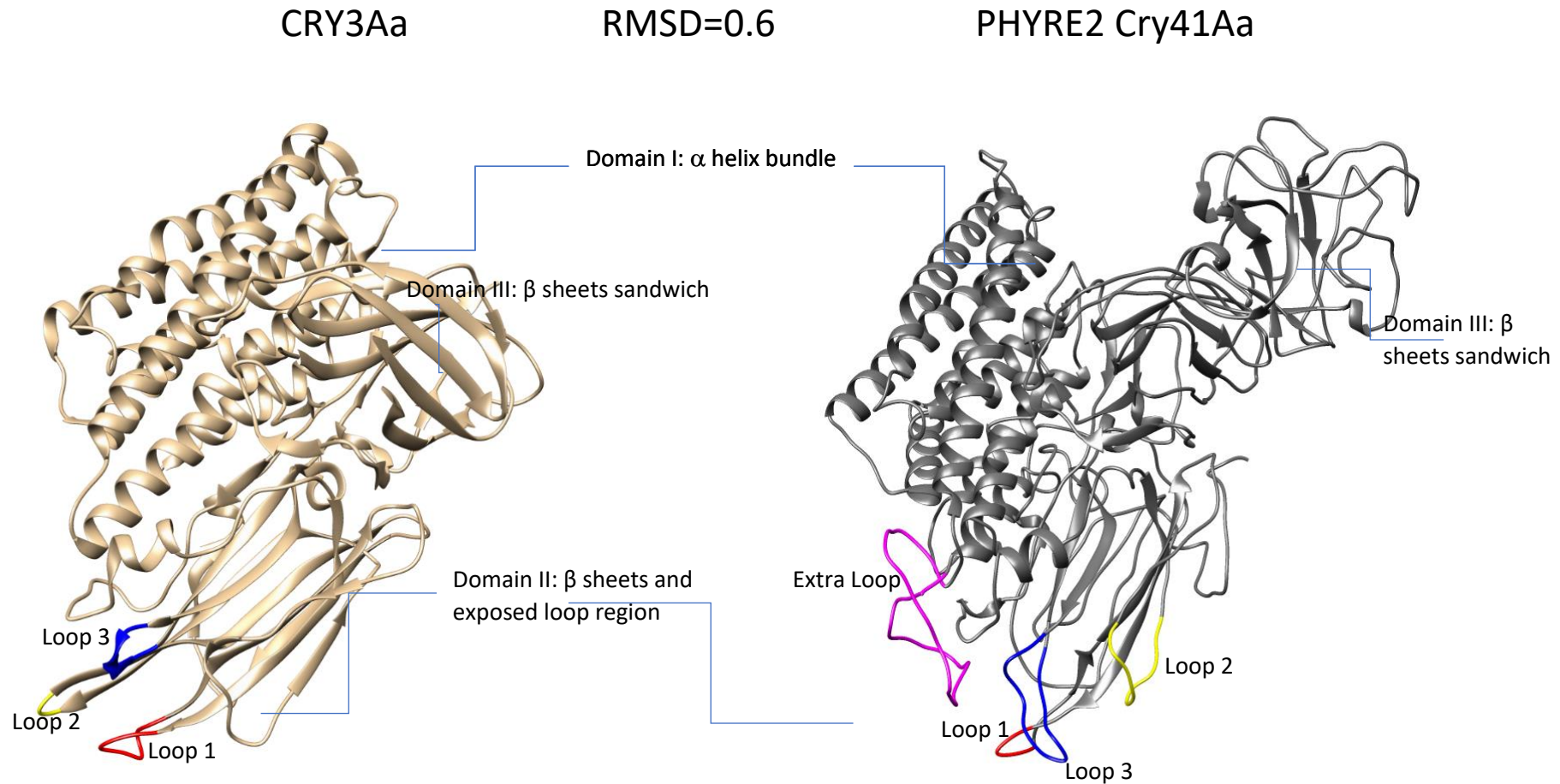
Figure 18 UCF chimeria visulaised Cry41Aa structure.

On the left is the 3D structure of crystallography resolved Cry3Aa shown in light grey. Loop1 is shown in red, loop2 in yellow, and loop 3 in blue. On thr right is the Phyre 2 generated predictaed strructure of Cry41Aa shown in dark grey. Loop1 is shown in red, loop2 in yellow, loop 3 in blue, and the extra loop exclusive to Cry41Aa in pin

The servers created similar but different models of Cry41Aa. The servers applied slightly different algorithms and programmes to analysis Cry41Aa sequence and resulted in each server's use of a different template. It worth analysing how different or similar the models are when each server applies the same template when analysing the Cry41Aa sequence. Thus, SWISS-MODEL analysed Cry41Aa sequence using Cry3Aa as template instead of Cry1Aa its highest scoring homologue. PHYRE2 analysed Cry41Aa sequence using Cry1Aa as a template instead of Cry3Aa. The two models were superimposed onto each relevant template and an RMSD value for each model was calculated as shown in figure 19.

SWISS-MODEL server created a new model of Cry41Aa. the most obvious change was observed in the extra loop that changed from its previous helical formation when Cry1Aa was template to a coil. The model and template were superimposed, and the RMSD was calculated as 1.5. The structure appears less compacted than the Cry41Aa model created with Cry1Aa as a template. The loops of domain II appear to project away from the main structure. The PHYRE2 server also created a new model of Cry41Aa based on the Cry1Aa template. The model and template were superimposed and the RMSD was calculated as 1.3. The model appears to be very similar to that made by SWISS-MODEL based on the same Cry1Aa template.

The RMSD values for each one of the four models made by both servers were compared. Both RMSD values for the given templates were above one (PHYRE2=1.3, SWISS-MODEL

=1.5), indicating that the models share little similarities to templates. The original templates chosen by the servers had RMSD scores closer to zero an indication that the similarities to those templates are greater. The lowest RMSD score was 0.6 by PHYRE2 server modelling Cry41Aa on the Cry3Aa template. This result indicates that this model shares the most similarities with its template and is thus the better model of Cry41Aa.

Figure 19 UCF chimeria visulaised structure of the two Cry41Aa models based on diffrent templates.

The superimposed models of Cry41Aa. In gold is the Cry41Aa model as predicted by SWISS_MODEL server based on Cry3Aa template. In light pink in the Cry41Aa model as predicted by PHYRE2 server based on Cry1Aa template . Loop1 is shown in red, loop2 in yellow, and loop 3 in blue and the extra loop exclusive to Cry41Aa in pink.

4.3 Discussion

Bioinformatic tools have highlighted similarities between Cry41Aa and the typical 3-domain insecticidal Cry toxins. It is thought to have a typical 3-domain structure as originally suggested by Yamashita *et al*. (2005).

Studies have highlighted two key points regarding the specificity of 3-domain Cry toxins. Firstly, domain II and its exposed loops are directly implicated in interactions with target receptors as well as being the region where most genetic variation is observed in 3-domain toxins. Secondly, the diversity that allows Cry toxins to have such a narrow specificity as is the case with Cry41Aa. research has suggested that receptors can recognise activated Cry toxins from more than one region of the protein. For example, monomeric Cry1Ab toxin is thought to bind to ALPs and APNs through loop 3 of Domain II and β16 of Domain III (Pacheco *et al*., 2009, Arenas *et al*., 2010). This is very different to the oligomeric form of Cry1Ab which interacts with ALP and APN through loop 2 of Domain II. In addition, studies have also suggested that Cry1Ab binds cadherin receptors through interactions with loop 2 and loop 3 and α8 of Domain II (Gomez *et al*., 2003, Gomez *et al*., 2006). Thus, it is important to explore regions associated with toxin specificity and the determinant of the toxin ability to interact with one cell type and not others.

The combined findings from the bioinformatics tools employed here, suggest that Cry41Aa has a 3-domain structure, with domain I, II and III maintaining the typical 3-domian arrangement of α and β sheets observed in insecticidal Cry toxins. ClustalW alignments and box shade analysis of Cry41Aa against protein sequences of resolved crystals structure of 3-domain Cry toxins have narrowed down the regions where the three loops of domain II Cry41Aa may occur. PSIPRED web server further analysed these regions predicting with a level of confidence to the location of β-sheets which typically fall either side of a loop.

Modelling tools in combination with the sequence analysis tools were able to model the putative exposed loops in the domain II region of Cry41Aa. Bioinformatics has suggested the following sequence for loop1 ^{384}SITS^{387}, loop2, ^{416}QNTSYTRIDRP^{426} and loop 3 ^{503}VRDNCPFAWPGYKQ^{L517} as the exposed loops of domain II in Cry41Aa. Homology modelling is a consistent trustworthy method to predict the 3-domain structure of a protein however it is not without errors; despite this even very inaccurate models are useful due to some aspects of function that is predicted from key structural features (Mahadeva Swamy *et al.*, 2014).

It has been argued that the diversity observed for some Cry toxins is due the variation of its domain II sequence. This domain has typically shown more variation compared with the other two domains. It is also argued that is this variation which is implicated in specificity and initial binding of Cry toxins (Dean *et al*., 1996, Pacheco *et al*., 2009,

Abdul-Rauf and Ellar, 1999, Pardo-Lopez *et al*., 2009, Abdullah and Dean, 2004, Lu *et al*., 1994). Therefore, it was not surprising that both protein modelling servers were able to model domain I well. Cry toxins share high amino acid sequence identity and similar tertiary structures (Lucena *et al.*, 2014). Previous research which used phylogenetic trees to study the domains of Cry toxins suggest that of the three domains, domain I is the most conserved. Genetic variation is likely to exist in the other two domains, particularly at the three exposed loops of domain II, which are thought to be critical sites for receptor recognition and binding. Thus genetic alterations in domain II loop region can dramatically affect the specificity of a 3-domain Cry toxin (Bravo *et al.*, 2018).

The search for Cry41Aa domain II loops began with a ClustralW alignment against resolved crystal Cry structures. This led to the predicted locations of loop 1, 2 and 3. It highlighted the distinction of the extra loop which did not align with any resolved Cry crystal structures. The smallest of loops, loop 1 aligned well with resolved Cry structures, followed by loop 3. Loop 2 was also predicted but it aligned the least with resolved Cry structures. PSIRED analysed the secondary structure of Cry41Aa. It indicated the location of β-sheets that fall either side of loop 1 and loop3, whose structures were predicted with high confidence. The secondary structure of loop 2 was predicted with low confidence and there were no immediate βsheets such as those observed for loop 1 and loop 3. Therefore, it was unclear where loop 2 began or where it ended. Due to this uncertainty, the study focused on the exploration of loop 1, 3 and the extra loop. The Cry41Aa was modelled by SWISS-MODEL and PHYRE2 servers. The

servers use different programmes and algorithms to build a model and therefore used different templates on which to base their Cry41Aa model. Both servers used insecticidal 3-domain Cry toxins as templates. However, in both cases sequence percentage identity was less than 37 %. The Cry41Aa models created by the servers indicated that Cry41Aa has 3-domain structure similar to insecticidal 3-domain Cry toxins. The extra loop was modelled to be helical by the SWISS_MODEL server whilst PHYRE2 modelled it as a coil. Domain I and III of both models appear very similar, in particular the α helical bundle of domain I. Loop 2 of the SWISS-MODEL Cry41Aa model has modelled the sequence as partially β sheet. This is supports findings of low confidence in the loop 2 structure.

The SWISS-MODEL server modelled Cry41Aa and matched it to Cry1Aa as its best homologue. When the model and Cry1Aa template were superimposed to calculate the RMSD it gave a value of 0.7. The closer the RMSD value is zero the more similar the two structures are. The SWISS-MODEL servers also calculated a GMQE score which is an estimate of the reliability of the alignment between a model and template. This model had a GMQE value of 0.48. The more reliable the estimate the closer the value to zero. Further analysis was carried out to calculate the QMEAN value of the model. This score indicates whether the model is comparable to experimental structures of similar size. The model had a QMEAN score of -4.0, which is consistent with a model of poor quality or low agreement with experimental structures of similar size. SWISS-MODEL server also modelled Cry41Aa based on Cry3Aa template. When the model was superimposed onto

the template the RMSD had a value of 1.5. This is indicative there is low similarity between this model and Cry3Aa.

The PHYRE 2 server modelled Cry41Aa where it covered 68 % of the sequence with 100 % confidence and modelled Cry41Aa to its nearest evolutionary fingerprint Cry3Aa. The model and template were superimposed and the RMSD was valued at 0.6. which is an improved RMSD score than that of the model by SWISS-MODEL server. The PHYRE 2 server modelled Cry41Aa on Cry1Aa the highest scoring homologue template used by SWISS-MODEL. It produces a model that did not differ greatly from that made by the SWISS-MODEL server. The model and Cry1Aa template were superimposed and the RMSD was valued at 1.3. RMSD value above one suggest low similarity between two structures.

In order to analyse the models by both servers it was important to standardise the values. There are no general standards for analysing models made by different protein modelling servers as the values are based on each servers' individual algorithms. However, a Q score is a value that considers the alignment length and the RMSD value. Q values above 0.4 indicate that the precision value for both the alignment and the RMSD are over 90 %. In order to analyse the models by both servers a Q score was calculated for each model. The SWISS-MODEL server modelled Cry41Aa with Cry1Aa as a template and had a Q score of 0.5, that is higher than the Q score of Cry41Aa with Cry3Aa as template which had a score of 0.3. The PHYRE2 server modelled Cry41Aa with

Cry3Aa as a template and had a Q score of 0.8. This is in contrast to the Q score of Cry41Aa model and Cry1Aa as a template which had the same score as the SWISS-MODEL best homologue template score of 0.5.

It is difficult to state which of the two servers provided the more accurate or more reliable model of Cry41Aa. The accuracy of a model can vary even within different regions of the same protein. It is usual that highly conserved regions are modelled with more accuracy than surface or loop regions. Generally, much of the error in modelling of unknown proteins comes from poor alignments with templates. Sequence similarities between a sequence and its template that have less than 40 % sequence identity tend to produce unreliable models. Both SWISS-MODEL and PHYRE2 have less than 37 % sequence identity with both Cry1Aa and Cry3Aa templates (Schwede *et al.,* 2003; Rother *et al.,* 2011).

Research on the specificity of 3-domain Cry toxin has applied mutagenesis techniques as an approach to gain insight into what makes a region or even an amino acid from a toxin affect the toxin's ability to interact with target receptors in such a way that it results in cell death.

Studies on loop 2 and loop 3 of Cry1Aa suggested that these loops play key role in their specificity and toxicity towards *Bombyx mori* larvae. Here, a mutation in Y445C of loop

3 reduced binding to BtR175 receptor and a deletion of N372 amino acid also reduced binding and resulted in a lower morality rate. However, when the same amino acid was substituted with an alanine residue the recombinant toxin exhibited enhanced toxicity to *Lymantria dispar in vitro* (Atsumi *et al*., 2008, Rajamohan *et al*., 1996a, Lucena *et al*., 2014). Further evidence from studies conducted on loop 3 of domain II of Cry1Aa has shown that substitutions between its native loop 3 with that of Cry4Aa resulted in an identical toxicity against *Culex pipiens* larvae (Abdullah *et al*., 2003). This approach may ultimately lead to the understanding of toxins specificity and shed light at least in part regarding its mode of action. This holds true particularly in the case of Cry41Aa where it has so far proved to have a narrow target cell (HL60 and HepG2) and is one of the least studied parasporins (Nagamatsu *et al*., 2010).

Such studies have highlighted the important role that domain II plays in the specificity of toxins and that identification and manipulation of domain II loops can introduce specificity as well as reduce or increase binding affinity and thus the toxicity of a toxin. It is clear that in order to learn about the specificity of Cry41Aa, its domain II loops need to be identified and explored. Yamashita *et al.* (2005) described Cry41Aa as a cancer killing Cry toxin. It is important to shed light on the regions which are responsible for this specificity. This information can be used to assist in the long-term fight against cancer. Furthermore, for the safe application of Cry toxins as bio-pesticides since it is known that they are able to target and interact with vertebrate cells lines.

5.0 Production of Cry41Aa hybrids and their activity

5.1 Introduction

The 3-domain Cry toxins have loops at the base of domain II (Schnepf *et al.*, 1998). In insecticidal Cry toxins, these are thought to recognise midgut receptors in target insects and bind irreversibly as a result of toxin insertion into the epithelia gut cell membrane resulting in subsequent pore formation and mortality (Pigott and Ellar, 2007; Smedley and Ellar, 1996; Rajamohan *et al.*, 1996; Adang *et al.,* 2014).

Data from resolved crystal structures of Cry toxins has been used to shed light on the function and role of domain II loops (Li, Carroll and Ellar, 1991). Information from resolved crystal structures of insecticidal Cry toxins has pointed to some of the approaches used to investigate Cry toxins and their mode of action and thus in part their specificity. These approaches include: (i) domain swapping between different Cry toxins that result in recombinant or chimeric proteins with hybrid domains, (ii) domain II loop exchange between different Cry toxins that result in recombinant proteins with domain II loops from a different Cry toxin.

This chapter details the mutagenesis carried out to create domain II loop exchange mutants of Cry41Aa and Cry41Ab with loops of insecticidal Cry1Ac as identified from bioinformatic analysis seen in chapter 4. It also details the mutagenesis carried out to create hybrids of Cry41Aa with insecticidal Cry42Aa and Cry1Ie. The mutagenesis approaches are summarised in table 7. The first approach involves an exchange of loop(s) between an insecticidal Cry toxin and Cry41Aa. The second approach involved the creation of domain II hybrids made from both insecticidal Cry toxins and Cry41Aa.

Mutagenesis approaches to study specificity of Cry41Aa

Mutagenesis approach	Toxin name	Site of mutagenesis
Loop exchange	Cry41Aa Cry41Ab Cry1Ac	Loop 3
Domain hybrids	Cry41Aa Cry42Aa Cry1Ie	Hybrid ORF2 of Cry41Aa &Cry42Aa Hybrid ORF2 withCry1Ie

Table 7 Mutagenesis approaches used in this study.
Table summarises the loop exchange and hybrid mutagenesis approaches employed to study Cry41Aa toxin specificity in this research.

All initial mutagenesis took place in *E. coli* expression vectors, these plasmids were approximately 5 Kb in size. They contained an antibiotic resistant gene as well as the ORF of the gene in question. Introduction and amplification of mutagenesis in these small plasmids resulted in less errors compared to introduction and amplification of mutagenesis in the much larger (~10 Kb) Bt expression vector.

The presence of a mutant in a plasmid capable of expressing only in *E. coli* meant that additional steps were required in order to introduce the mutant in a plasmid that would express in a *Bt* host. Mutant (insert) DNA and *Bt* backbone fragments were isolated with a double enzyme digest and allowed to ligate in order to form a *Bt* expression vector capable of mutant protein expression in a *Bt* host.

At each stage Hae III digest profiles and sequencing results confirmed the correct sequence of the mutant. This also meant that the mutant required confirmation after its introduction into a *Bt* shuttle vector. It was important to confirm that the data collected on any expressed recombinant proteins was due to the intended mutagenesis. Thus, the plasmids were extracted from the *Bt* host. However, due to the presence of native *Bt* plasmids it was not possible to carry out Hae III directly on *Bt* extracted plasmids. A Hae III digest profile here would result in too many bands, rendering the profile illegible. Further *E. coli* transformations were carried out before Hae III digests/sequencing confirmed the correct sequence of mutagenesis in ORF of Cry gene from plasmids extracted out of *Bt* host. An overview of the key steps and check points of the mutagenesis are detailed in figure 20.
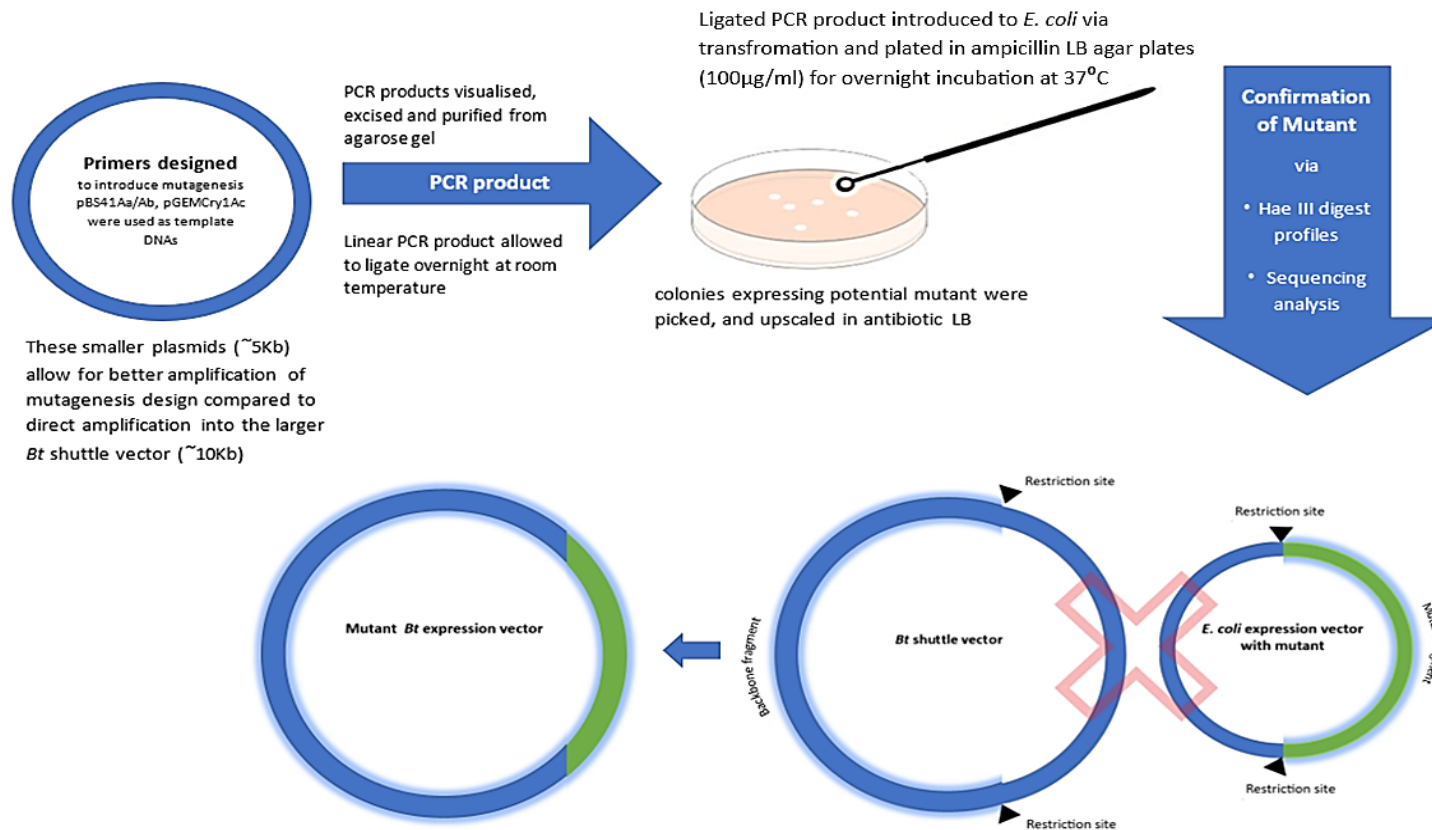
Figure 20 Mutagenesis steps to create a mutant. An overview of the initial steps taken to create a mutant plasmid that can be expressed in a *Bt* host cell. The process begins with a PCR design and results in a mutant *Bt* shuttle vector.

For the purpose of this study, plasmids that acted as templates or where mutagenesis has been designed to take place, are referred to as wildtype. This is despite that fact that these plasmids are not strictly wildtype and do not exist in the natural genome of the *Bt* strain on which they originate from. Plasmids that have had alterations in their gene(s) because of the mutagenesis design are referred to as constructs.

Cry41Aa and Cry41Ab were expressed from ORF2 and ORF3 genes. First, ORF2 of Cry41Aa was cloned into pBluescript SK2+ plasmid to create wildtype pBS41Aa. ORF2 of wildtype Cry41Aa is flanked by a BamHI 5' and an XhoI enzyme restriction site at the 3' end. It contains an ampicillin resistance gene as well as an *E. coli* origin of replication as shown in figure 21. Similarly, Cry41Ab was also cloned into pBluescript SK2 to create wildtype pBS41Ab also shown in figure 21. All mutagenesis was initially carried out using wildtype pBS41Aa and pBS41Ab which acted as template DNA. Other plasmids were also used to create mutants of Cry41Aa and Cry41Ab these were introduced accordingly with each mutagenesis approach.
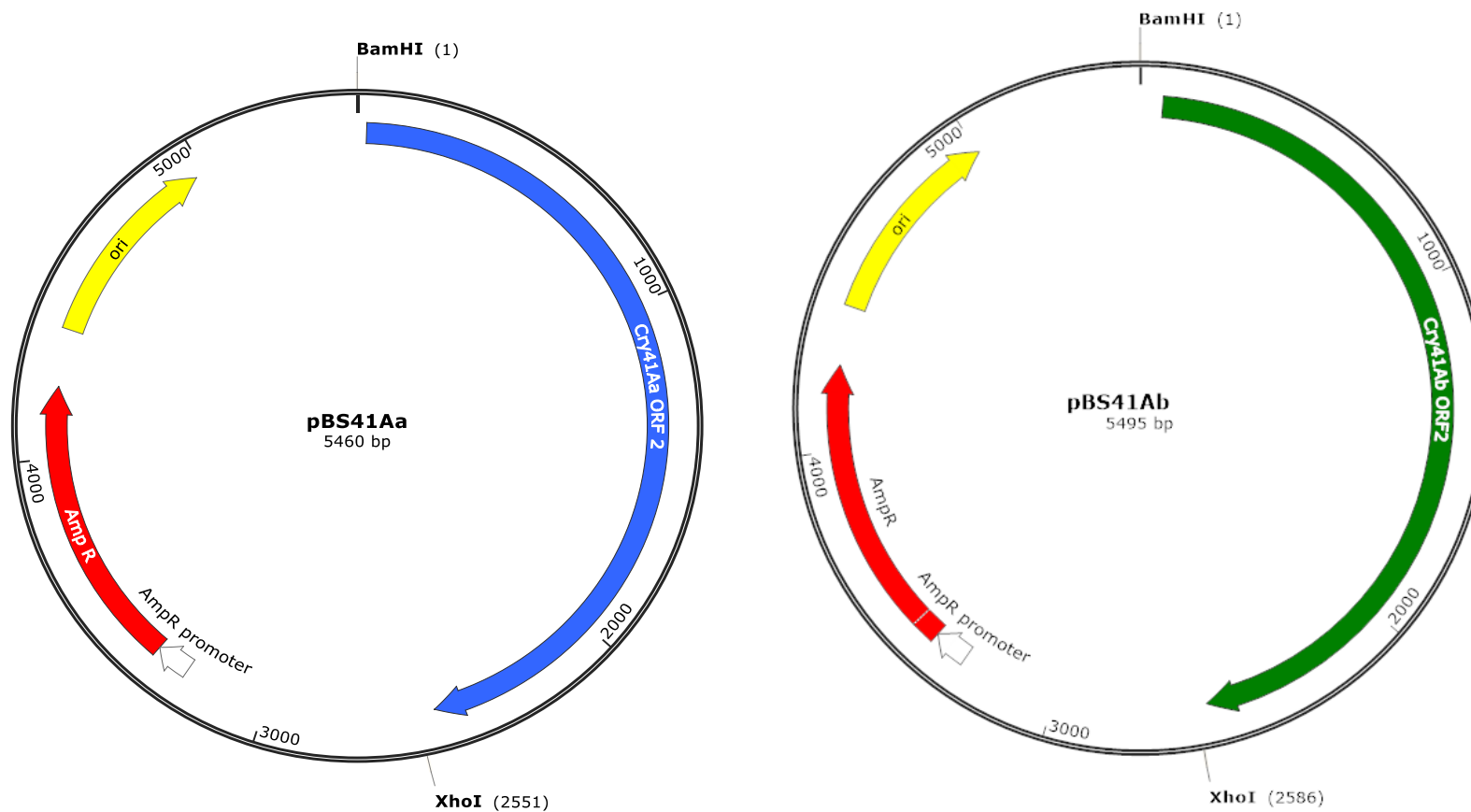
Figure 21 Schematic diagram of plasmids involved in insecticidal and cytocidal loop 3 exchange.
 On the left wildtype plasmid pBS41Aa with ORF2 of Cry41Aa shown in blue. On the right is wildtype pBS41Ab with ORF2 of Cry41Ab shown in dark green. Ampicillin resistance genes shown in red and bacterial origin of replication in yellow. BamHI and XhoI restriction sites flank either side of OFR2.
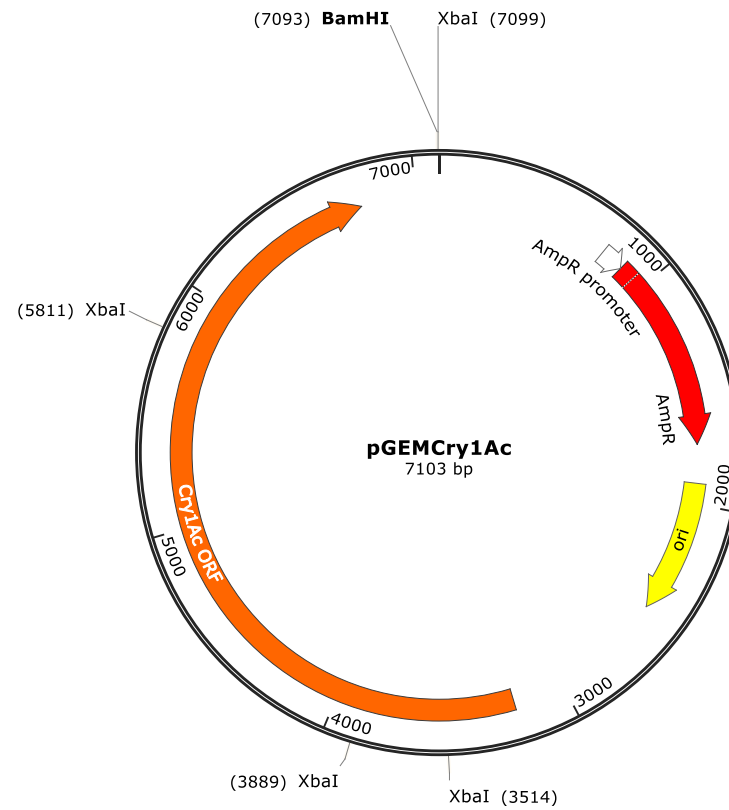
Figure 22 Schematic diagram of pGemCry1Ac plasmid
Wildtype pGemCry1Ac plasmid, with Cry1Ac ORF in orange, and an ampicillin resistance gene shown in red. Bacterial origin of replication shown in yellow. Restriction sites for BamHI and XbaI shown.

5.2 Cry41Aa loop exchange mutagenesis in loop 3 of domain II

Mutagenesis via loop exchange has been employed to explore Cry toxin specificity. Several studies have engineered *Bt* Cry toxins by exchanging loop regions of domain II and thus introducing a new specificity to the recipient Cry toxin. (Abdullah *et al*., 2003, Liu and Dean, 2006, Howlader *et al*., 2009, Pardo-Lopez *et al*., 2009).

Cry toxins differ in their primary sequences and target specificities. The 3-domain structure of Cry toxins have a similar structure but can differ in their primary sequence and have different cell target specificities (Crickmore *et al*., 1998, Bravo *et al*., 2011). Their shared structural similarity has been taken to explore the three loops at the apex of domain II, and studies that have modified this region have resulted in novel specificities, potent toxicity, and changes to the stability of a toxin (Bravo *et al*., 2011, de Maag *et al*., 2001; Lucena *et al*., 2014; Pigott and Ellar., 2007; Bravo *et al*., 2011; Bravo *et al*., 2013; Nachimuthu and Polumetla Ananda., 2004; Florez *et al*., 2012) .

In a study aimed to determine the role of domain II loops in the specificity of Cry1Ah and Cry1Ai*,* a loop exchange investigation was carried out. Cry1Ah is toxic to *Helicoverpa armigera* but not toxic to *Bombyx mori* larvae. The closely related Cry1Ai is toxic against *B. mori* but is not active against *H. armigera*. A loop 2 and 3 exchange between these Cry toxins resulted in hybrid toxins with an enhanced toxicity, confirming that it is loops2 and 3 of Cry1Ah that are responsible for receptor binding and toxicity to *H. armigera* (Zhou *et al*., 2017).

This part of the study aimed to introduce a new insecticidal specificity to Cry41Aa in an attempt to locate the regions responsible for its specificity towards the HepG2 cell line.

Cry1Ac is an insecticidal 3-domain Cry toxin of ~130 KDa and has toxicity towards *Manduca sexta* and other lepidopteran insects (Alouf, Ladant and Popoff, 2015; Schnepf *et al*., 1998; de Maagd *et al*., 1999; de Maagd, Bravo and Crickmore, 2001; Crickmore *et al*., 1998). This insecticidal Cry toxin is not known to be toxic to vertebrate cell lines or HepG2. It has a typical 3-domain structure and the wildtype pGEMCry1Ac plasmid was readily available at the time of this study. Here, Cry1Ac acted as the negative control for Cry41Aa's target cells HepG2. The mutagenesis design aimed to remove the loop 3 of both cytocidal Cry toxins (Cry41Aa and Cry41Ab) and replace it with loop 3 of insecticidal Cry1Ac. If successfully expressed, the loop exchange mutants would be investigated for potential newly acquired specificities that they did not possess before, by testing on them on various mammalian cell lines. In other words, whether cytocidal Cry41Aa has acquired a new specificity as a consequence of the introduction of loop 3 from Cry1Ac and vice versa.

The wildtype plasmid pGEMCry1Ac acts as template DNA in a PCR reaction designed to introduce loop 3 of Cry41Aa and Cry41Ab into Cry1Ac and replace its own loop 3. The wild type pGEMCry1Ac is shown in figure 22. It contains a single ORF with conserved blocks 1 to 8 and has a BamHI 5' site and three XbaI restriction sites together with an ampicillin resistant gene. The wildtype loop 3 of Cry1Ac is made up of seven residues

[428]SGFSNSS[433]. The primers were designed to amplify the entire plasmid except for loop 3 of Cry1Ac, here the seven residues are replaced by loop 3 of Cry41Aa which is made up of fifteen residues [503]VRDNCPFAWPGYKQL[517]. Figure 23 details the Primer design for the PCR reaction that introduced the loop 3 exchange mutagenesis approach. The new sequences are incorporated into the 5' end of primers. Table 8 lists the primers used to create the constructs.

Loop 3 primer list.

Template DNA	Primer name	Oligonucleotide/ Primer sequence	5'-PHO
pBS41Aa	41Ac L3F	5'AGT AGT GTA AGT AGT CGT TTG TTA TTT GGT TG 3'	Yes
	4111AcL3R	5'GTT ACT AAA GCC TGA AGG TTC AAA TTT AAT CCA AG 3'	Yes
	41Ac L3F	5'AGT AGT GTA AGT AGT CGT TTG TTA TTT GGT TG 3'	Yes
pBS41Ab	41AcL3R	5'GTT ACT AAA GCC TGA AGG TTC AAA TTT AAT CCA AG 3'	Yes
	1Ac41L3F	5'TGGCCTGGTTATAAACAATTATAAGAGCTCCTATGTTCTC3'	Yes
pGEMCry1Ac	1Ac41L3R	5'GGCGAAAGGGCAATTATCCCGTACACGAAACATTGAAACATGG3'	Yes

Table 8 Loop 3 primer list.
Table lists the primers used to create loop 3 exchange constructs. The sequence belonging to Cry41Aa is underlined in green. The sequence belonging to cry1Ac is underlined in orange.
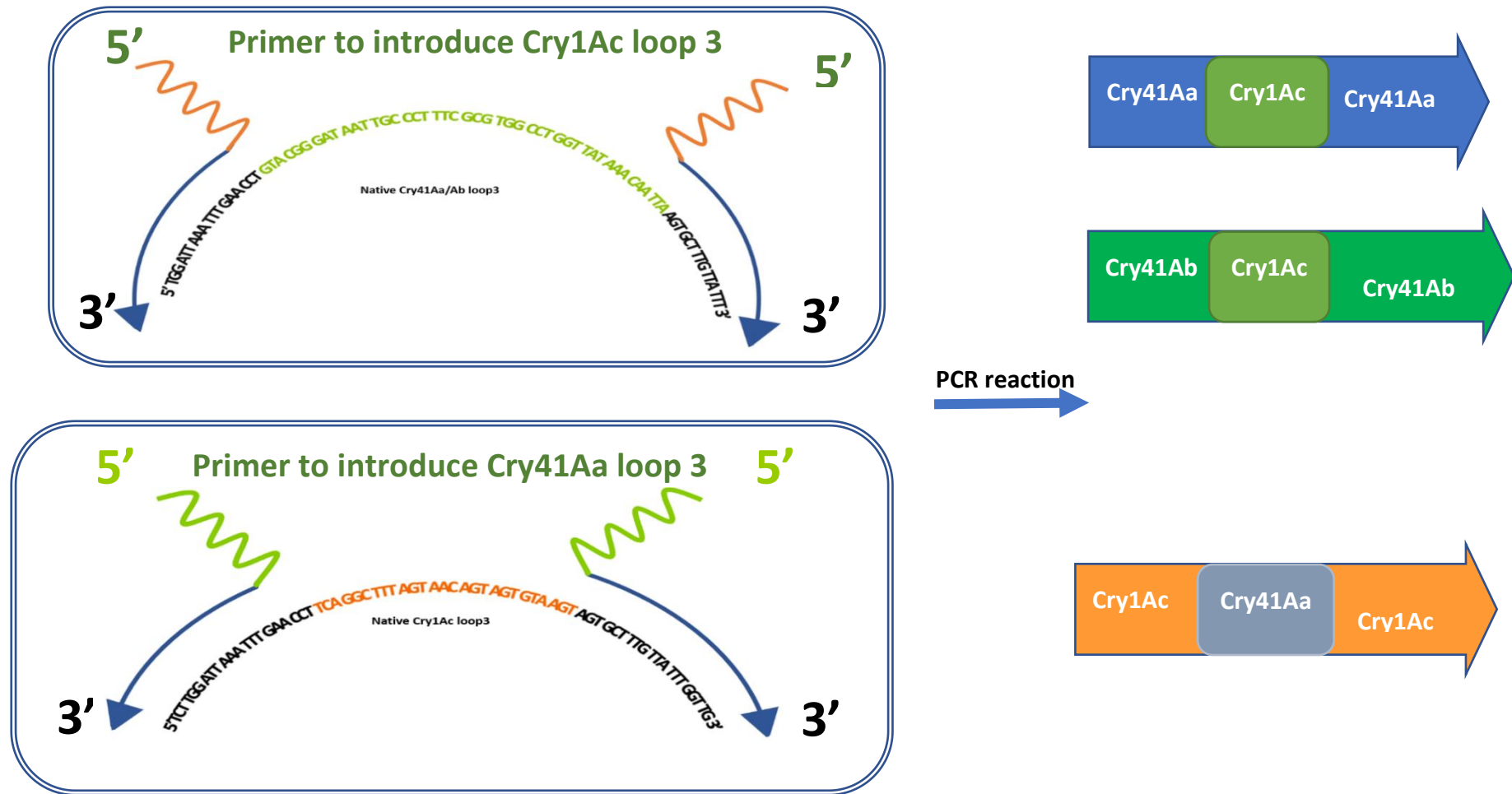
Figure 23 PCR primer design for loop 3 substitution.
ORF2 of Cry41Aa is shown in blue with loop 3 of Cry1Ac in orange. ORF3 of Cry41Ab is shown in green with loop 3 of Cry1AC in orange. ORF of Cry1Ac shown in orange with loop 3 of Cry41Aa in blue.

The resulting linear DNA was gel purified, allowed to ligate overnight and introduced into *E. coli* and plated onto ampicillin LB agar plates. Potential colonies were picked and allowed to grow in ampicillin LB broth or re-streaked on ampicillin agar plates. A mini prep was carried out to isolate plasmids with potentially correct loop 3 exchange construct. To check for correct configuration of the construct, Hae III digests were carried out on both the wildtype pGEMCry1Ac plasmid and the pGEMCry1Ac41 loop 3 constructs. This construct has the ORF of insecticidal Cry1Ac with loop 3 of Cry41Aa. The resulting Hae III digest profile of pGEM1Ac41loop3 was analysed by NEB cutter online software. The NEB cutter predicated 21 fragments as a result of a Hae III carried out in correct construct pGEMCry1Ac41loop3. It further predicted 20 fragments as a result of a Hae III carried out in the wildtype pGEMCry1Ac plasmid. Figure 24 shows the picture of the agarose gel analysis of the both Hae III digests. Table 9 lists the key fragments for both digests, which were identified in the gel picture of the Hae III digest of both wildtype and construct plasmid. The construct pGEMCry1Ac41loop3 has two distinct Hae III digest fragments 1035 bp and the 348 bp. The 289 and 285 bp bands appear as a doublet as shown in figure 24 and their absence in the wildtype pGEMCry1Ac lane is clear. The construct was then confirmed by sequencing analysis.

Hae III digest profile by *NEB web* cutter

Hae III fragments	Wildtype pGEMCry1Ac (bp)	pGEMCry1Ac41loop3 Construct (bp)
1	1542	1542
2	1365	1035
3	908	908
4	654	654
5	434	458
6	289	434
7	285	348
8	279	289
9	267	285
10	174	279

Table 9 Hae III digest profile by *NEB web* cutter
Table lists the key fragements of the Hae III   digest profile for wiltype pGEMCry1Ac and the construct pGEMCry1Ac41loop3.
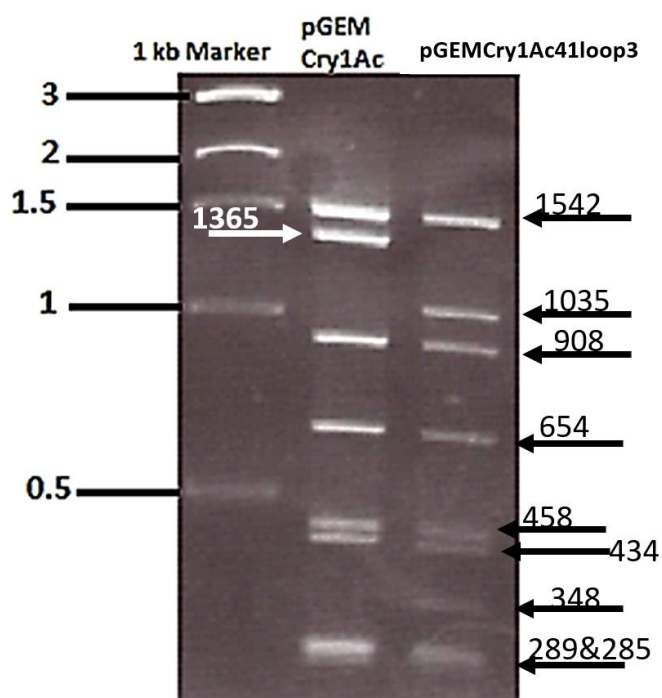

Figure 24 Banding profie of construct in agrose gel
Agarose 1.5% gel showing DNA marker in lane 1. Hae III digest banding profile for wildtype pGEMCry1Ac in lane 2 and hybrid construct in lane 3.

Once confirmed, the construct pGEMCry1Ac was incubated in ampicillin LB for 3 days at 37 ^{o}C to express and produce the loop 3 exchange hybrid pGEMCry1Ac41loop3. The cells were then harvested and sonicated to collect the hybrid toxin for further analysis. Figure 25 is the SDS PAGE gel picture of the crude hybrid crystals of pGEMCry1Ac41loop3. The crude crystals were characterised in a similar process as conditions used to characterise other 3-domain Cry toxins. The samples were solubilised in carbonate buffer pH 10.5 and treated with trypsin 1 mg/mL. The resulting solubilised trypsin activated hybrid protein was visualised in SDS PAGE gel as shown in Figure 26. This trypsin stable hybrid was analysed in HepG2 cell assay to assess whether it had gained cytocidal activity as a consequence of the replacing of its wildtype loop with that of Cry41Aa.
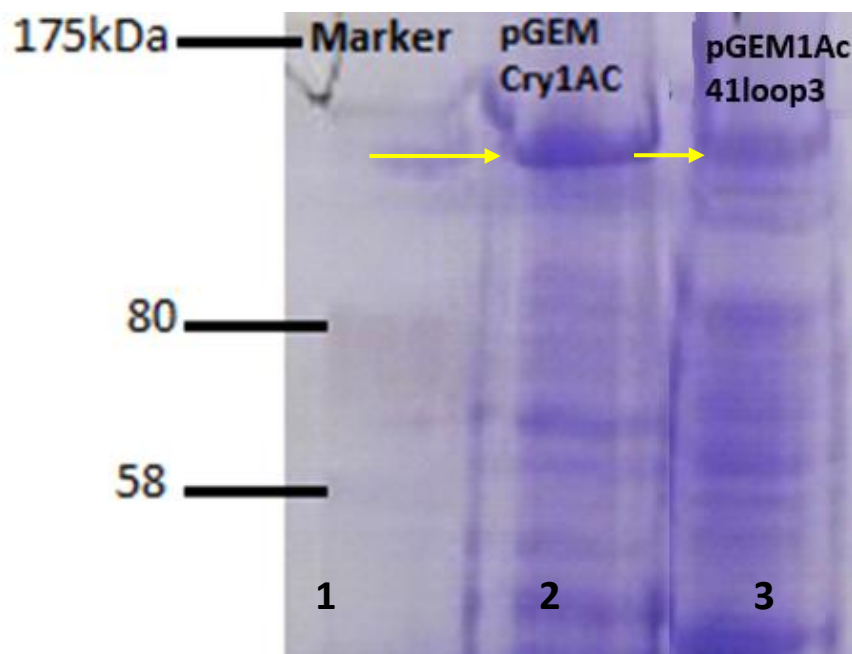

Figure 25 SDS-PAGE analysis showing crude sample for hybrid crystals.
Lane 1 contained protein marker. Lane 2 contained wildtype Cry1Ac crude crystals harvested from *E. coli* cells yellow arrow indicated ˜130 KDa band. Lane 3 contained crude crystals harvested from *E. coli* cells containing the hybrid plasmid pGEMCry1Ac41loop3, yellow arrow indicating ˜130 KDa band. 5µL of parasporal inclusion protein samples were loaded in each lane.
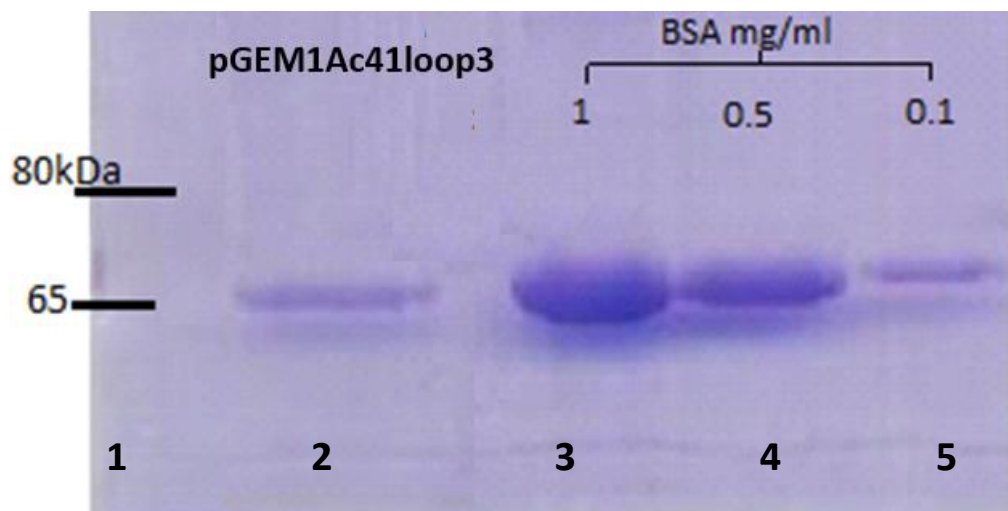
Figure 26 SDS-PAGE gel showing 65 KDa solubilised trypsin treated Cry1Ac41loop3.
Lane 1 contained protein marker. Lane 2 contained a 65 KDa solubilised trypsin treated hybrid protein expressed from pGEM1Acloop3 plasmid in *E. coli.* Lanes 3, 4, and 5 contained. BSA concentrations range 1 mg/mL, 0.5 and 0.1 respectively.

Preliminary cell assays were carried out on the recombinant pGEM1Ac41loop3. The recombinant was incubated with HepG2 cell lines for 24 h , thereafter the cell viability was measured using CellTiter-Blue assay as shown in figure 27. It approximates the metabolic activity of cells using a fluorometric method. The resazurin dye is nontoxic but permeable to cells. Once it enters viable cells, it is reduced to high fluorescent resorufin and the signal is measured to give an estimate of viable cells (O'Brien *et al*., 2000). Insecticidal wildtype Cry1Ac acted as a negative control and did not demonstrated any toxicity towards HepG2 cells. Cry1Ca acted as a control for 3-domain crystal proteins made in host *Bt4D7*. Wildtype Cry41Aa demonstrated toxicity to the HepG2 cell. It showed similar toxicity to etoposide which acted a positive control. Triton -X100 is a

detergent that permeabilises cell membranes and here it also acted as positive control.

The protein concentrations were not optimised. The recombinant protein made by

construct pGEM1Ac41loop3 did not gain toxicity to HepG2 cells.
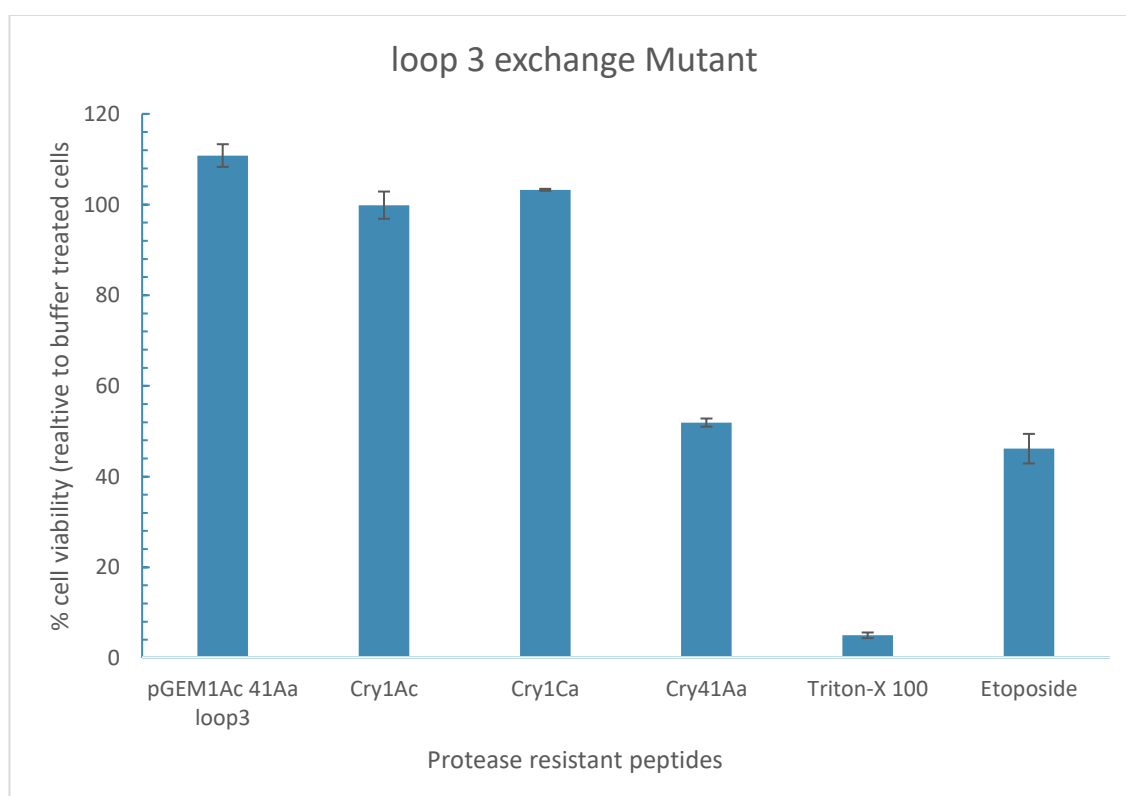

Figure 27 24hr CellTiter-Blue HepG2 cell viability assay of loop 3 mutant
Graph of the fluorescent signal measurements from HepG2 cell line 24 h after toxin exposure. Cry1Ca act as negative controls. TX-100 and Etoposide act as positive controls. Approximately 100µg /mL of protein was added.

Wildtype pBS41Aa and pBS41Ab were used as DNA templates for primers designed to

remove native loop 3 from their domain II and introduce loop 3 of insecticidal Cry1Ac in

its place via PCR as detailed in the loop exchange primer design, figure 23. The PCR

products were gel purified and allowed to ligate over night at room temperature before

transformation into *E. coli* cells. The mixtures were plated onto ampicillin LB agar plates and were incubated overnight. Colonies with potential constructs pBS41Aaloop3 and pBS41Abloop3 were picked and allowed to incubate in ampicillin LB broth or re-streaked onto ampicillin LB agar plates for overnight incubation. A mini prep was carried out to extract construct plasmids from *E. coli* cells. Hae III digests were carried out to authenticate the correct constructs. The construct pBS41Aaloop3 was digested by the Hae III enzyme. The NEB web cutter predicted the fragment sizes, this digest profile is shown in table 10. Figure 28 shows the agarose gel analysis of pSB41Aaloop3 construct. It has the distinctive 684 bp fragment absent in the Hae III digest profile of wildtype pBS41Aa plasmid (wildtype not shown in the gel image).

Similarly, pBS41Abloop3 construct was subjected to a Hae III digest. Table 11 lists the predicted Hae III digest profile for wildtype pBA41Ab and the pBA41Abloop construct. As suggested from the predicted fragment sizes it would be difficult to distinguish the wildtype apart from the construct. Figure 29 shows the agarose analysis of Hae III digest performed on plasmid mini prepped from colonies 2b, 1a, 1c, and 1e alongside the Hae III digest of wildtype pBS41Aa that acts as reference, but not a comparison, to construct pBA41Abloop3. Both constructs were later further confirmed by sequencing protocol.

Hae III digest profile by NEB web cutter.

Hae III fragments	Wildtype pBS41Aa (bp)	pBS41Aaloop3 Construct (bp)
1	767	767
2	629	684
3	558	629
4	549	549
5	458	458
6	434	434
7	340	340
8	306	306
9	267	267
10	254	254

Table 10 Hae III digest profile by *NEB web* cutter.
Table lists the key fragements of the Hae III digest profile for wiltype pBS41Aa and the construct pBS414a1loop3
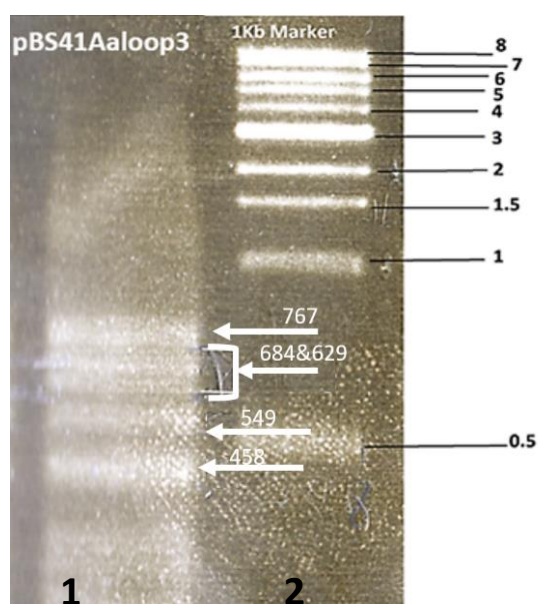

Figure 28 banding profie of construct in agrose gel
Agarose 1.5% gel showing Hae III digest banding profile for pBS41Aaloop3 in lane 1. DNA marker in lane 2

Hae III digest profile by NEB web cutter

Hae III fragments	Wildtype pBS41Ab (bp)	pBS41Abloop3 Construct (bp)
1	1001	1106
2	767	767
3	629	629
4	458	458
5	436	436
6	434	434
7	378	378
8	267	267
9	254	254
10	174	174

Table 11 Hae III digest profile by NEB web cutter
lists the key fragments of the Hae III digest profile for wiltype pBS41Aa and the construct pBS414b1loop3
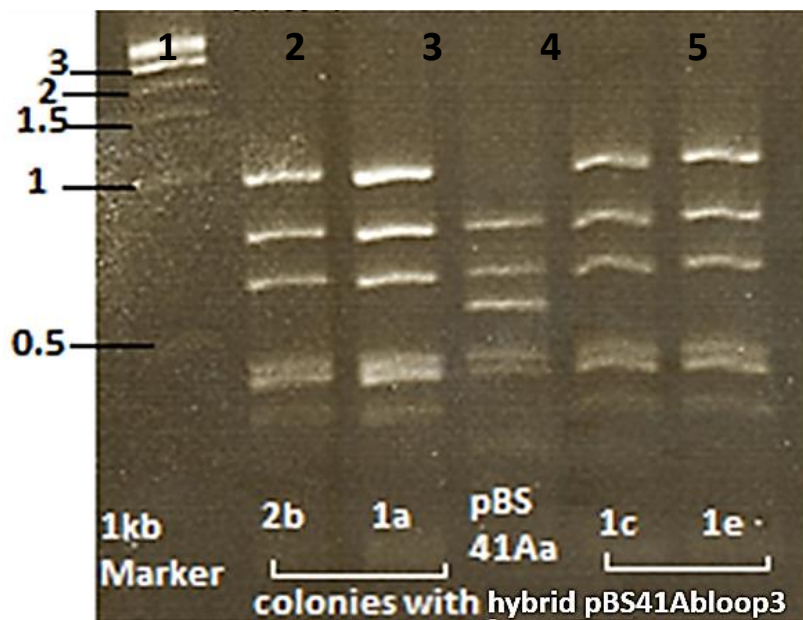

Figure 29 Banding profie of construct in agrose gel.
Agarose 1.5% gel showing Hae III digest banding profile for pBS41Aaloop3 in lane 2, 3, 5,6, control pBS41Aa in lane 4. DNA marker in lane 1

Both, pBS41Aaloop3 and pBS41Abloop3 construct plasmids lack a *Bt* origin of replication and cannot be expressed in Cry41Aa cells. Furthermore, the constructs lack the ORF3 gene associated with crystal formation in Cry41Aa. Therefore, the loop 3 constructs required subcloning into a vector with a *Bt* origin of replication and Cry41Aa ORF3 gene in order to express the recombinant *Bt* crystal. For the purpose of this study, a *Bt* shuttle vector was used and referred to as wildtype pSVP2741Aa plasmid. pSVP2741Aa is a *Bt* expression vector ˜10 Kb in size.

To create it, both OFR2 and ORF3 were cloned into the *Bt* shuttle vector pSVP27. The ORF2 gene was flanked by two restriction sites, a BamHI and XhoI site on either side of the gene. In the same manner the OFR2 in pBS41Aa/Ab was also flanked by the same restriction sites for BamHI and XhoI enzymes. This allows for the subcloning of the mutant OFR2 from blue script plasmid into the *Bt* shuttle vector. Previous attempts to perform mutagenesis directly on the *Bt expression* vector pSVP2741Aa have proven difficult due to its large size (Krishnan., 2013). The ORF3 of Cry41Aa and Cry41Ab have 99% sequence homology (Yamashita *et al.*, 2005), for the purposed of this study the same ORF3 from Cry41Aa was used to express potential hybrids of both Cry41Aa and Cry41Ab. Figure 30 shows the schematic representation of wildtype pSVP2741Aa and pSVP2741Ab. Both *Bt* expression vectors have a strong *Bt* Cyt1Aa promoter that drives ORF2 and ORF3 expression in the acrystalliferous *Bt* 4D7 host strain. Furthermore, they contain ampicillin and chloramphenicol resistant genes that allow for selection for colonies of interest with potential correct constructs.
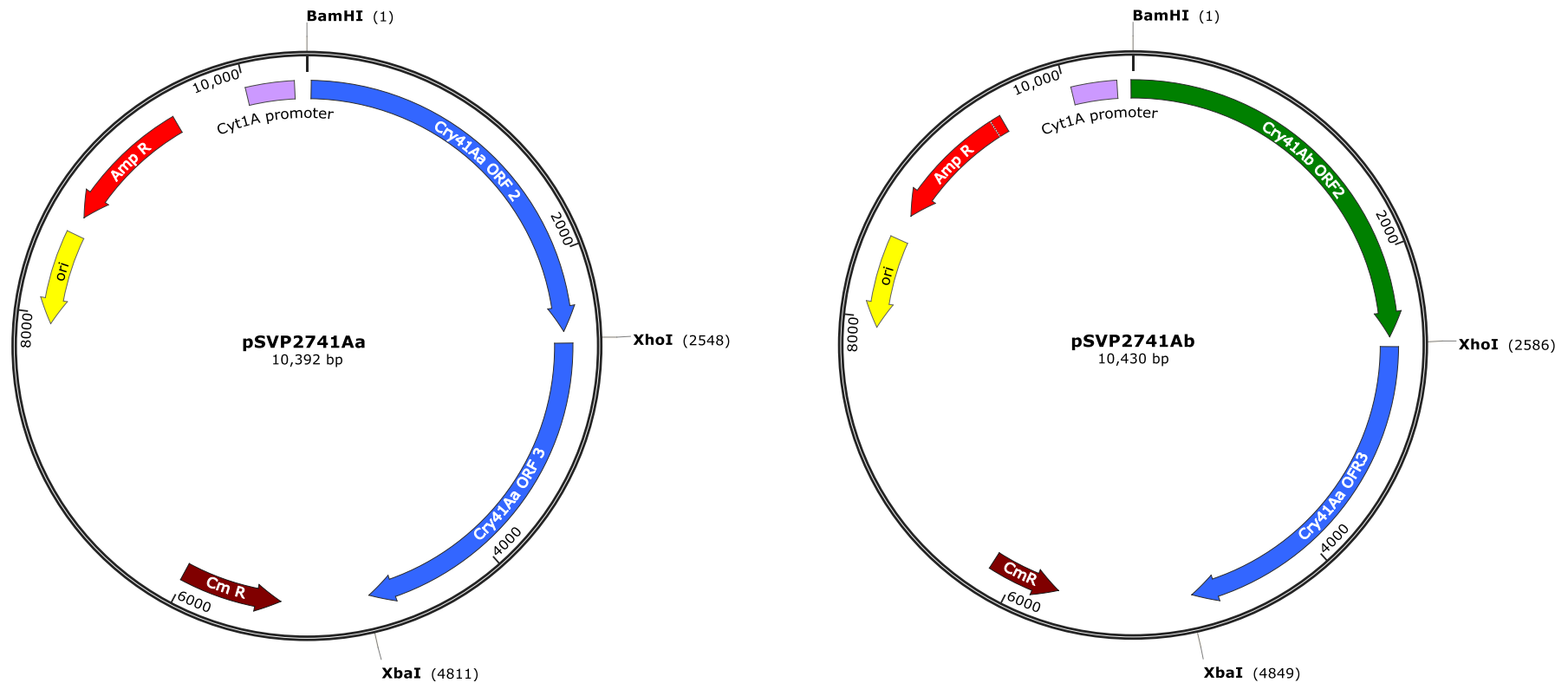
Figure 30 Schematic presentation of wildtype pSVP2741Aa and pSVP2741Ab. pSVP2741Aa (left) and pSVP2741Ab (right) with ampicillin resistant gene (red) and chloramphenicol resistance gene (brown), as well as bacterial origin of replication. ORF2 of Cry41Aa shown in blue in pSVP2741Aa. ORF3 of pSVP2741Ab shown in green. ORF3 of Cry41Aa shown in both plasmids in blue.

The next step of the loop 3 mutagenesis was to introduce the mutated ORF2 from constructs pBS41Aaloop3 and pBS41Abloop3 into the larger pSVP2741Aa/Ab expression vector. The restriction enzymes BamHI and XhoI allow for subcloning of ORF2 into the pSVP2741Aa *Bt* expression vector.

The constructs pBS41Aaloop3 and pBS41Abloop3 underwent a double digest by BamHI and XhoI. The linear DNA was run on an agarose gel and the 2.5Kb linear DNAs encoding loop 3 exchange ORF2 were cut out and the gel purified in preparation for ligation. Similarly, the wildtype pSVP2741Aa plasmid also underwent a double digest by BamHI and XhoI enzymes. The linear DNAs were also run on an agarose gel, and the 7.8 Kb linear DNA encoding OFR3 was cut out and the gel purified in preparation for ligation. The 2.5Kb linear DNAs from constructs were allowed to ligate overnight with 7.8kb backbone DNAs from *Bt* expression vector pSVP2741Aa. When correctly ligated these fragments form pSVP2741AaXhoI1Acloop3 (OFR2 of Cry41Aa with loop 3 of Cry1Ac) and pSVP2741AbXhoI1Acloop3 (ORF2 of Cry41Ab with loop 3 of Cry1Ac) constructs.

E. coli cells were transformed with ligation mixtures *and* plated out on LB agar plates prepared with both 5 µg/mL chloramphenicol and 100 µg/mL ampicillin for overnight incubation at 37 °C. Colonies harbouring the potential constructs were picked and incubated in antibiotic LB or streaked onto antibiotic LB agar plates.

Mimi preps were carried out to extract potential constructs from *E. coli* cells. These were subjected to Hae III digest to check for the correct loop 3 exchange constructs in *Bt* expression vectors.

Table 12 lists the Hae III digest profiles for wildtype pSVP2741Aa and both loop 3 exchange constructs. The agarose gel analysis of both constructs is shown in figure 31 and 32. Figure 31 is the agarose gel image of the Hae III digest of construct pSVP2741AaXhoI1Acloop3, here the distinct 684bp band is seen when compared to wildtype. This construct was further confirmed by sequencing.

Figure 32 is the agarose gel image of the Hae III digest of construct pSVP2741AbXhoI1Acloop3 which has the distinct 1155bp band absent in wildtype pSVP2741Aa. The construct was further confirmed by sequencing. Once confirmed, both constructs were introduced into *E. coli* strain *GM*2163, which removes methyl groups from the DNA and facilitates transformation of the Bt 4D7 host and then plated out on antibiotic agar plates.

The cells were scraped and mini prepped to extract the construct plasmids. Construct plasmids pSVP2741AaXhoI1Acloop3 and pSVP2741AbXhoI1Acloop3 were now ready for *Bt4D7* transformation. The transformation mixtures were plated onto 5 µg/mL chloramphenicol prepared LB agar plates and were incubated overnight at 30°C.

Transformations with *Bt* are notoriously difficult and it took numerous attempts to introduce the *Bt* expression constructs to *Bt* cells.

The colonies harbouring potential constructs were picked and incubated in 2-3mL of LB antibiotic broth for 2-3 h. The mixture was divided for the following usage. First it was poured and spread out in 5 µg/mL chloramphenicol plates and incubated for 3 days at 30 ^{o}C. The remaining mixture was centrifuged, and pellet was miniprepped to extract plasmids from *Bt* cells.

Hae III digest profile by NEB web cutter

Hae III fragments	Wildtype pSVP2741Aa (bp)	pSVP2741AaXhoI1Acloop3 Construct (bp)	pSVP2741AbXhoI1Acloop3 Construct (bp)
1	2469	2469	2469
2	1895	1895	1895
3	1007	1007	1155
4	879	879	1045
5	629	684	879
6	598	629	629
7	587	598	587
8	558	587	458
9	458	458	436
10	434	434	434
11	306	306	267

Table 12 Hae III digest profile by NEB web cutter
Table lists the key fragements of the Hae III digest profile for wiltype pSVP2741Aa   and the constructs pSVP2741AaXhoI1Acloop3 and pSVP2741AbXhoI1Ac1loop3.
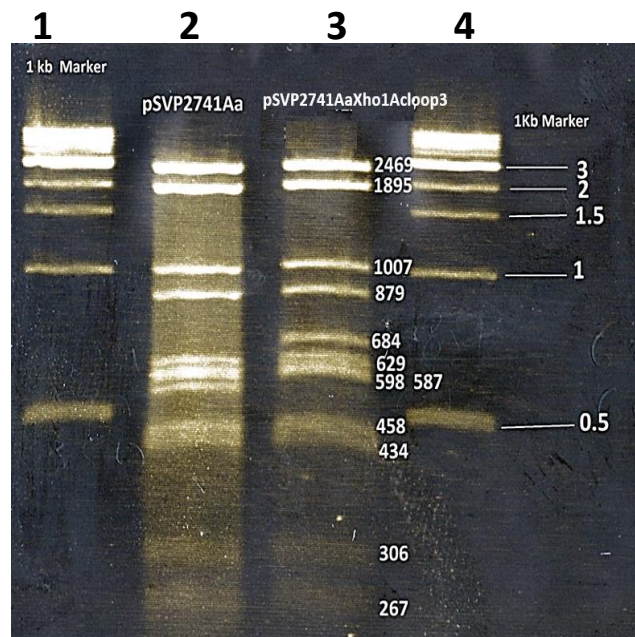
Figure 31 banding profie of construct in agrose gel
Agarose 1.5% gel showing Hae III fragment sizes of hybrid pSVP2741AaXho1Acloop3. Lane 1 contained DNA marker, lane 2 wildtype pSVP2741Aa, lane 3 had hybrid pSVP2741AaXho1Acloop3. Lane 4 contained DNA marker.
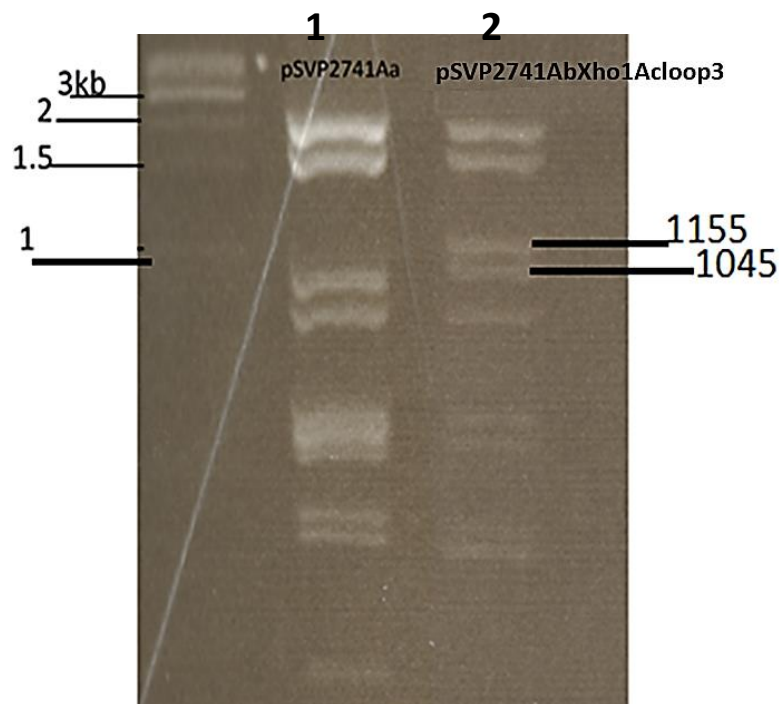

Figure 32 banding profie of construct in agrose gel
Agarose 1.5% gel image of Hae III digest of pSVP2741AbXho1Acloop3 hybrid. Lane 1 contained DNA marker, lane 2 wildtype pSVP2741Aa and lane 3 contained pSVP2741AbXho1Acloop3 hybrid

A Hae III digest was not be performed directly on *Bt* extracted plasmids as the mini prep mixture included both native *Bt* plasmids and the plasmid of interest. In order to eliminate native *Bt* plasmids from the mini prep mixture, a final transformation of *E. coli* cells was carried out and plated out on LB agar plates prepared with both 5 μg/mL chloramphenicol and 100 μg/mL ampicillin for overnight incubation at 37°C. Here the plate should only have *E. coli* cells that harbour the construct plasmids only. The bacterial lawn was scraped, and a mini prep was performed to extract the plasmids. A Hae III digest was carried out on these plasmids to confirm the integrity of the construct after *Bt* cells transformation.

After 3 days, sample plates were analysed under a light microscope to check for the presence of crystals and spores. The presence of crystals and spores are signs that *Bt* cells have undergone the sporulation stage as a result of limited nutrition (Crickmore *et al*., 1998, de Maagd *et al*., 2003). Once confirmed the cells were harvested and sonicated to obtain hybrid crystals for characterisation. Crude crystals of loop 3 exchange hybrids were analysed in SDS PAGE gels and compared to wildtype Cry41Aa and Cry41Ab crude samples, as shown in figure 33. SDS PAGE analysis indicated there is no obvious difference between wildtype Cry41Aa/Ab crystals and their loop 3 exchange hybrids.
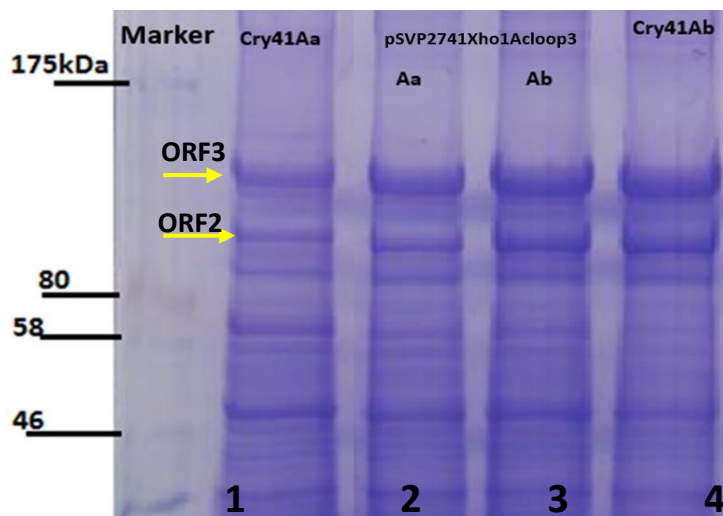
Figure 33: SDS-PAGE gel showing crude samples of loop 3 exchange
Crude samples of loop 3 exchange hybrid crystals from constructs pSVP2741Aa XhoI1Acloop and pSVP2741Ab XhoI1Acloop. Lane contained crude Cry41Aa. Lane 2 crude recombinant pSVP2741Xho1Acloop3Aa. crude recombinant pSVP2741Xho1Acloop3Ab. Lane 4 contained crude Cry41Ab. Samples in lanes labelled respectively, each lane contained 5µL of parasporal inclusion protein. ORF2 is a ˜88 KDa band is indicated with yellow arrows. ORF3 is ˜120 KDa band indicated with yellow arrow.

All crystals were solubilised in carbonate buffer pH 10.5 and activated with trypsin (1 µg/mL). Despite changing the conditions, it was not possible to activate loop 3 exchange. Proteins from the hybrids of Cry41Aa/Ab with loop 3 of Cry1Ac were not stable and degraded in the presence of the trypsin. This suggests that hybrids were structurally unstable due to the loop exchange which affected the stability of the hybrid.

5.3 Cry41Aa ORF2 hybrids with other Cry genes

In previous studies Cry toxin hybrids have been constructed in attempts to learn more about their specificity and exploit their toxicity (Crickmore *et al.*, 1998; Dean *et al.*, 1996; de Maagd *et al.*, 2003). It usually involves the mutagenesis of loops by either deletion or

substitution of residues (Adang, *et al.*, 2014). Domain hybrids created by the deletion, addition, and or exchange of domains have played a key part in the improvement of known Cry bio pesticides as well as the creation of novel toxins. The removal, addition and exchange of Cry toxin domains has revealed information about the regions responsible for toxin specificity and toxicity.

In a study to improve toxicity towards sap sucking insects (Hemiptera), a 12 amino acid sequence peptide known to bind to the epithelial gut of pea aphid insects was added to the loops of the cytolytic toxin Cyt2Aa. This addition resulted in the improved binding and toxicity towards *Acyrthosiphon pisum (*pea aphid) and *Myzus persicae (*peach aphid) (Chougule *et al.*, 2013; Bravo *et al.*, 2013).

Interactions of structural domains I, II, and III were investigated by the creation of chimeric proteins from domain I and III combinations of Cry1Ab, Cry1Ac, Cry1C, and Cry1E. The effects of the chimeric toxins were tested on Sf9 cells, a clonal isolate of *S. frugiperda* Sf21 cells commonly used in insect cell culture. Of the parental Cry toxins, only Cry1C was known to affect cell viability and membrane permeability. The study found that chimeric proteins with domain II of Cry1C showed activity against cells. Further investigation indicated that domain II from an active toxin is required for toxin stability but did not result in an active toxin capable of cell death. Pore size caused by chimeric toxins in the Sf9 cells membrane varied depending on which domain I was present in the chimeric toxin. The study suggested that pore properties are a direct result of interactions between domain I and domain II or III. It further speculates that it

is domain III which regulates the level of toxicity exhorted by chimeric toxins when a chimeric toxin made up of domain III from Cry1Ab and domain II and III from Cry1C was more toxic than the paternal Cry1C toxin (Rang *et al*., 1999).

In a study to improve the biological control of pest insects *Tecia solanivora* (*Lepidoptera: Gelechiidae*), and *Hypothenemus hampei* (*Coleoptera: Scolytidae*), a hybrid toxin from toxic Cry strains was created by domain swapping between Cry1Ba and Cry1Ia. The resulting chimeric toxin had an improved toxicity when compared to the parental Cry toxin strains (López-Pazos *et al*., 2010) . In an attempt to apply this approach, Cry41Aa hybrids with Cry42Aa and Cry1Ie were designed.

5.3.1 Cry41Aa ORF2 hybrid with Cry42Aa

The gene for *Cry42Aa* was previously identified from the genome of *Bacillus thuringensis A1462* (Krishnan, 2013). Little is known about Cry42Aa, except that it is very similar to Cry41Aa. It is a split toxin with two open reading frames, whereby ORF2 produces a 70 KDa peptide and ORF3 produces a 120 KDa peptide.

It is not clear whether the expression of Cry*42Aa* was silenced or if the protein produced from its expression led to an inactive toxin. Previous experiments on Cry*42Aa* have shown that it solubilises in alkaline conditions. However, despite various attempts to activate the protoxin of Cry42Aa it was not possible to obtain a protease resistant core.

In an attempt to learn more about the specificity of Cry41Aa and also produce a stable active toxin from Cry42Aa a hybrid of Cry41Aa and Cry42Aa was designed as follows.

A multiple sequence alignment was carried between the protein sequences of *Cry41Aa* and *Cry42Aa* as shown in figure 34.

The construct design aimed to have a full length ORF2 made up of both *cry41A* and *cry42Aa*. This is to ensure that the five conserved of Cry toxins present in both *cry41Aa* and *cry42Aa* are complete and propagate protein folding of the hybrid crystal.

Höfte and Whiteley, (1989) first described the five conserved blocks of 3-domain Cry toxins as sequences that encompass the Cry protein active core of domain I, II, and III. Schnepf *et al.* (1998) later described the remaining conserved blocks (Höfte and Whiteley, 1989; Crickmore *et al.*, 1998). Conserved block 1 is situated in the central helix of domain I, block 2 is found at the domain I and domain II interface, and block 3 is located at the boundary between domains II and domain III. Block 4 is in the central β-strand of domain III and block 5 is found at the end of domain III (Xu *et al.*, 2014; Palma *et al.*, 2014).

```
CLUSTAL 2.1 multiple sequence alignment


Cry41Aa     MNQNCNNNGYEVLNSGKGYCQPRYPFAQAPGSELQNMGYKEWMNMCTSGDPTVLGEG--Y
Cry42Aa     MNQNYNNNGLEILDSG-GVCSPRYPLANAPGSELQNMGYKEWLEMCSIKGAETFADKSTL
            **** **** *:*.** * *.****:*:************:.**:  .. .:.:

Cry41Aa     SADVRDAVITSINIASYLLSVPFPPAGVAAGILGALLGLLWPTNTQAVWEAFMNTVEALI
Cry42Aa     SAQSQEGLRTAITIALSLLSNLPGPFGYPAKLLSIIFPFLWPTNTQAQWEAFMKVVEELV
            **: ::.: *:*.**  ***    * * .* :*. :: :******** *****:.** *:

Cry41Aa     NQKLDEYARSKAISELNGLKNVLELYQDAADDWNENPGDLRNKNRVLTEFRNVNGHFENS
Cry42Aa     DQKIETFARDQAIQRLRGIQDVISLYQRDAKNFNDYPTSEPIQRQLLSQFTATNTFIVGS
            :**:: :**.:**..*.*:::*:.***  *.::*: * .   :.::*::*  .* .: .*

Cry41Aa     MPSFAVRNFEVNLLPVYAEAANLHLLLLRDAVKFGEGWGMSTDPGAERDDMYRRLRSRTE
Cry42Aa     MSLFRVGRHEVPLLTTFVQAANLHLLLLRDAIMFGESWGMCPVTVAGYQNDFN---NRIA
            *. * *  ..** **..:.:***********: ***.***.. .   *   :: ::   .*
*. * * ..** **...::*************: ***.***..
                                                Extra Loop

Cry41Aa     IYTDHCVNTYNQGLQQAKSLQANVSDYSRYPWTQYNQSGGFSYREAKGEYRGTENWNLYN
Cry42Aa     DYTDYSVSIYNQGLQKAKTLKANLRDYEKYPWARYYN----SSVGPEFAYGDMENWNLYN
             ***:.*. ******:**:*:**. **.:***::** :    *   .: * . *******

Cry41Aa     AFRRDMTILVLDIIAQFPTYDPGLYSRPVKSELTREVYTDIRGTTWRSDANLNTIDAIEN
Cry42Aa     NYRRDMTLMVLDLVALWPTYNPQQYPIAPKIQLTREIYTELRGNAG--NTKRPSMDAIDA
             :*****::***:** :***:* *.    :.:::  **: .  :*****::***::* :***:*
*.  .  * :*****:**::**.:   :::   ::***:
                                    Loop1

Cry41Aa     RMVGSRQLQLFTWLTEMKFYIRNTGSITSYTHGDLMVGLEKKIRKTNDNDQWLPLEGQNT
Cry42Aa     ELIPPP--RLFTWLESVDMHRWPT-SAGYYYYTFQNAGIKHRYKYTLDS-QTLTSSLRGA
            .::  .     :*****  .:.::      * *    * :     .*::::  :  * *.  *
*.  .  .:.:              .::  .   :*****  .:.::    * *   * :     .*::::  :  * *.  *
*.  .  .:.:
            Loop2
Cry41Aa     SYTRIDRPGIELGKNYWYYARTQQWFETRLLQLWANTDVLSLNAGTVGNEFWVRDVPDYR
Cry42Aa     SGNNFLVPAEETINRVQNQHGEGLYTFSFYRSGQ--SDPFLNIGTTADKPYVSTMNRIP
            *  ..:::   *   *      :: : :   ::       .   ** **..:: :*   :
                                Loop3
Cry41Aa     NIYARSTRNHFIENHRLSWIKFEPVRDNCPFAWPGY--KQLSALLFGWTHNSVDLNNIIS
Cry42Aa     -----VEGDQTQANHRLSWITGMVIPELSIPAFGHYNPTYISCAAEGWTHLSVERSNEIK
                 ::   *******.  :  : .  *:  *   .:*.    ****  **: .*  *.

Cry41Aa     QYRITQIPAVKAYWNRGAFSVIRGPGSTGGNLVQLGTGG----EVSVKVRPEQTGS-DWY
Cry42Aa     SDKITQIPAVKAFQLSNNASVVRGPGSTGGDLVQFSATSSGNKQLWIKVKPTTIALGRRF
            . :**********:     .  **:********:***.:.     ::  :**:*    .   :

Cry41Aa     RVRIRYAAGSRGRLNVKKYVSSIHASVTYDYNMTMSSSTQGTYNSFQYLDVYNFRLAEPE
Cry42Aa     KVRIRYAAAANVTFTVQKCVTGVACWETATKSVTTTYSGTLTYNAFKYVDIFEIPANESE
            :*******::   :.*:* *:.: .  *    .:* : *    ***:*:*:*::::    *.*

Cry41Aa     FEVWLTNESGGPIWIDKIEFIPLSPIPELPVYPGTYQIVTALNNSSVVTSEEFCMGIGLT
Cry42Aa     FSLEFLSTSGGPIYIDKIEFIPVNPIPEPPVPEGIYQIVTALNNSSVVDMDPGTWGT---
            *.: : .  *****:*********:. ***** * * ************* :    *
                                           Ricin domain
Cry41Aa     TRCGVNLWSNNGNTLQKWRFVYNGDQNAFQIKSTPNEDLVLSGSNSGTSVTAETNQNRPN
Cry42Aa     -RHNVHLWQNNNTNNQKWRFVYNSSQGAYQIRNLADENLVLTREGANVKVVSYQNNN-TA
             * .*:**.**...  ******.. *.*:*:**:. .:*:***: ......*.:  *:*. .

Cry41Aa     QYWLIEEAGNGYVYLRSKGNPNLVLDVAGTSTANGTNIILWNYNGSTNQKFKLSMNYNVT
Cry42Aa     QYWIIEDAGNEYVYLKSKADPSRVLDVTGSSTQNGTNIQVWSNYGTLNQKFKLVKL----
            ***:**:*** ****:** .:*. ****:*.** ***** :*. *: ****** ******

Cry41Aa     KAREAVQALFSNPTTLQLKVTDHHVNQVARLVECIADQIHPKEKMCLLDQVKLAKRLSRE
```

```
Cry42Aa           ------------------------------------------------------------

Cry41Aa           RNLLNYGDFESSDWVGTDGWNVSTNVYTVADNPIFKDHYLNMPSANNPILSDKIFPTYAY
Cry42Aa           ------------------------------------------------------------

Cry41Aa           QKVEESRLKPYTRYIVRGFVGSSKDLEILVARYDKEVHKRMNVPNDIIPTSPCTGEPVSQ
Cry42Aa           ------------------------------------------------------------

Cry41Aa           PTPYPVMPSNTMPQDMWCNPCGNGYQTAAGMMVQSTGMMCQDPHEFKFHIDIGELDMERN
Cry42Aa           ------------------------------------------------------------

Cry41Aa           LGIWIGFKVGTTEGMATLDNIEVVEVGPLTGDALTRMQKRETKWKQKLTEKRMKIEKAVQ
Cry42Aa           ------------------------------------------------------------

Cry41Aa           IARDAIQTLFTCPNQSCLQSAITLQNILRAEKLVQKIPYVYNQFLQGVLSAVPGEAYAYD
Cry42Aa           ------------------------------------------------------------

Cry41Aa           IFQQLSDAVATARALYNQRNVLNNGDFSAGLSNWNGTEGADVQQIGNASVLVISDWSASL
Cry42Aa           ------------------------------------------------------------

Cry41Aa           SQHVYVKPEHSYLLRVTARKEGSGEGYVTISDGTEENTETLKFMVGEETTGATMSTIRSN
Cry42Aa           ------------------------------------------------------------

Cry41Aa           IRERYNERNMATPDPDAYGGTNGYASNQNMVNYSSENYGMSAHSGNNNMNYQSESFGSKP
Cry42Aa           ------------------------------------------------------------

Cry41Aa           YGDGNSMINGSSNNYEANGYPGNNNINDQSENYGANAYSSNNMNYQSESSGFTPYGDENN
Cry42Aa           ------------------------------------------------------------

Cry41Aa           MTNYPSNNYEMNPYSSDMNMSMNRGSDCGCGCSANAYPGGNMMMNNYSSSTYEMNTYPSS
Cry42Aa           ------------------------------------------------------------

Cry41Aa           TNMTNHQGMGCGCHYSTNEYPMIEENIPDFSGYVTKTVEIFPETNRVCIEIGETAGTFMV
Cry42Aa           ------------------------------------------------------------

Cry41Aa           ESIELIRMDCE
Cry42Aa           -----------
```

Figure 34 Alignment of Cry41Aa and Cry42Aa

The clustal W alignment of the protein sequences of Cry41Aa and Cry42Aa. The conserved blocks of both proteins are underlined and in colour font as follows: conserved block 1 in red, conserved block 2 in orange, conserved block 3 in green, conserved block 4 in pink, conserved block 5 in blue, (only Cry41Aa) conserved block 6 in grey, conserved block 7 in purple, conserved block 8 in brown. Domain II of Cry41Aa are highlighted as follows: Loop1 in red, loop 2 in yellow, loop 3 in navy and extra loop in pink. The ricin domain is highlighted in teal.

The five conserved blocks of 3-domain toxins are usually found in the proteolytic active core post activation. Domain I is associated with pore formation and is comprised of conserved block 1 and part of conserved block 2. Domain II is associated with specificity; it is comprised in part of conserved block 2 and conserved block 3. Domain III is associated with binding and receptor recognition. It is usually comprised in part of conserved block 3, and the whole of conserved block 4 and 5 (Palm a *et al*., 2014).

The construct was designed to contain conserved blocks 1 and 2 from Cry42Aa and conserved blocks 3, 4, and 5 from Cry41Aa. Cry42Aa and Cry41Aa share homology in the domain I region. Introducing this homologous region from Cry42Aa into Cry41Aa will indicate if this region of Cry42Aa is functionally active. Conserved block 3, 4, and 5 from Cry41Aa, where less homology between the Cry42Aa and Cry41Aa is observed, ensures the presence of regions associated with specificity in insecticidal 3-domain toxins. Their inclusion also aims to maintain in part the structural integrity of Cry41Aa.

The hybrid was initially made by using primers to amplify conserved block region for blocks 3, 4, and 5 from the pBS41Aa plasmid (figure 21) used as template DNA in chapter 5.1 In addition to this, the construct also had an ampicillin resistant gene and an *E. coli* origin of replication. Similarly, the conserved blocks 1 and 2 of Cry42Aa were amplified using primers for the DNA fragment from pSVP2742Aa template DNA. The primers are listed in the table 13 below. The two linear DNAs were run on an agarose, cut out and the gel was purified. The fragments were allowed to ligate overnight before an *E. coli.*

Transformation. The ligations formed a compete ORF2 that was made up in part from

Cry42Aa (block1 and 2) and Cry41Aa (blocks 3, 4, and 5).

Figure 35 is the schematic representation of the pBS41Aa_Cry42Aa construct. ORF2 is a

hybrid gene that is flanked on either side by BamHI and XhoI restriction sites and allows

for subcloning of the hybrid gene into a *Bt* shuttle vector at a later stage. The

transformation mixture was incubated overnight on 100 µg/mL ampicillin prepared LB

plates at 37°C. If correctly ligated, it would create the pSVP2741Aa_Cry42Aa construct.

Colonies that potentially harboured the construct were picked and incubated in

antibiotic LB broth for 2-3 h. The samples were centrifuged and resuspended the pellet

underwent a mini prep to extract plasmids from the bacterial cells.

PCR primer list to create construct

Template DNA	Primer name	Oligonucleotide/ Primer sequence	5'-PHO
pBS41Aa	CP41AaF	5' GATCCGGGACTTTATAGTAGGC3'	No
	CP41AaR	5'TGACAATCCTCCATTCCATTTG3'	No
pSVP2741_Nhe	CPCry42F	5'ATGAATCAAAATTATAACAACAATGG3'	Yes
	CPCry42R	5'ATATGTTGGCCATAACGCAAC3'	Yes

Table 13 PCR primer list to create construct
 lists the PCR primers to create hybrid construct Cry41-42Aa
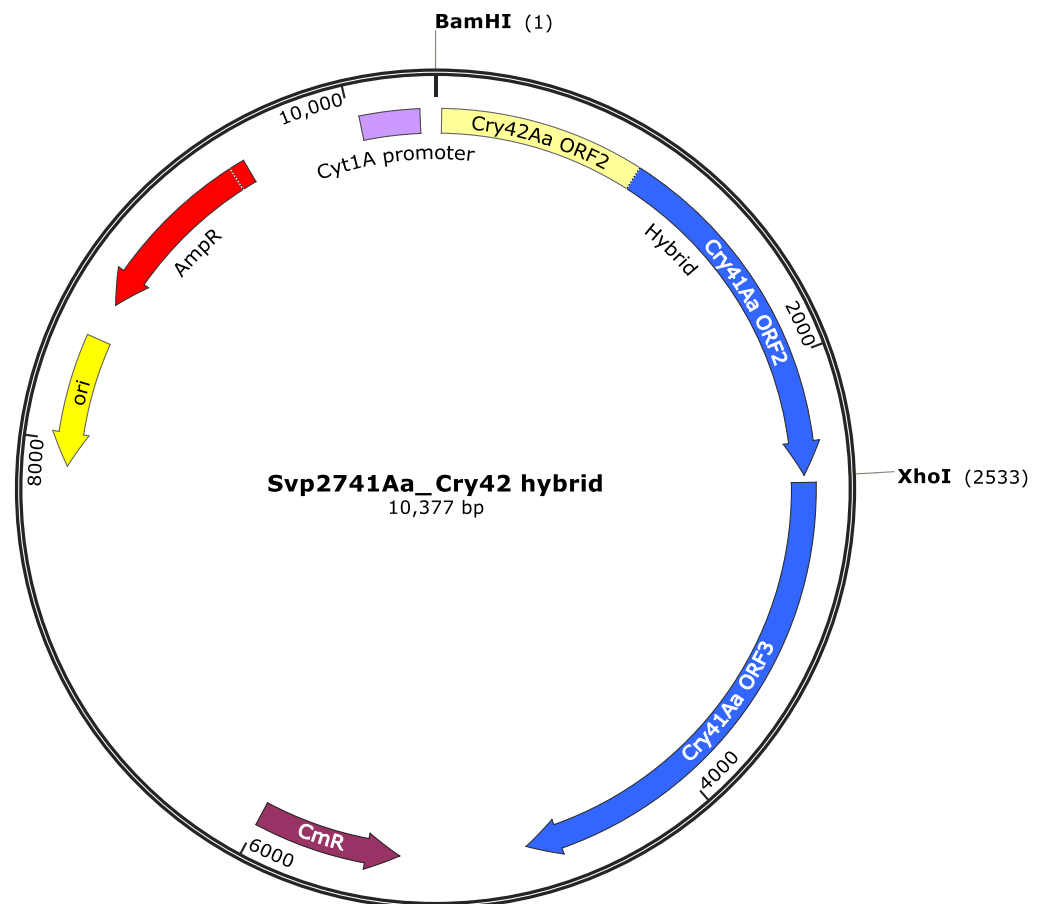
Figure 37 schematic presentations of hybrid construct.
 pSVP2741Aa_Cry42Aa with ampicillin resistant gene (red) and chloramphenicol resistance gene (brown), as well as bacterial origin of replication. OFR2 is hybrid gene made up of both Cry41Aa and Cry42Aa. ORF3 of Cry41Aa is shown in blue.

Colonies with potential construct were picked and incubated in antibiotic LB broth for 2-3 h. The mixture was centrifuged, and pellet resuspended for extraction of plasmids. A mini prep was performed to extract potential constructs from bacterial cells. These were subjected to Hae III digests and analysed in agarose gel. The NEB web cutter website has predicted the fragment sizes for both the wildtype pSVP2741Aa and correctly ligated construct pSVP2741Aa_Cry42Aa. The information is detailed in table 15.

Figure 38 is the gel picture of the different Hae III carried out at different stages of the pSVP2741Aa_Cry42Aa construct. First the construct was confirmed after Hae III was performed in plasmids extracted from *E. coli JM*109. The distinctive 959bp and 759bp bands were observed.

Once sequencing data had confirmed the construct, it was introduced into *E. coli GM*2163 in preparation for transformation of *Bt4D7*. *Bt4D7* transformants were plated out on 5 μg/mL chloramphenicol LB agar plates and incubated overnight 30°C. *Bt* colonies were picked and incubated in 3mLs of antibiotic LB broth for 2-3 h. The mixture was divided, 2mLs was distributed between 5 μg/mL chloramphenicol LB agar plates for a 3-day incubation period at 30°C. The remaining 1 mL was centrifuged down, and the pellet was resuspended in water to extract the construct from *Bt* cells.

Mini preps carried out in *Bt4D7* would simultaneously extract native *Bt* plasmids. Hae III performed on such a sample would produce many fragments making the digest profile for the construct difficult to confirm. Therefore, the miniprep from *Bt* cells was introduced into *E. coli JM109* and incubated overnight in 5 µg/mL chloramphenicol and 100 µg/mL ampicillin LB agar prepared plates. The lawn should only have the construct pSVP2741Aa_Cry42Aa. The cells were scraped off and miniprepped to prepare the extracted plasmids with a Hae III digest. The outcome of the Bt profile of the pSVP2741Aa_Cry42Aa construct was consistent and confirmed the integrity of the construct. After 3-day incubations samples of the *Bt* agar plates were analysed under the light microscope for crystals and spores.

Hae III digest profile by *NEB web* cutter

Hae III fragments	Wildtype (bp)	pSVP2741Aa	pSVP2741Aa_Cry42Aa Construct (bp)
1	2469		2469
2	1895		1895
3	1007		959
4	879		879
5	629		601
6	598		587
7	587		558
8	558		458
9	458		434
10	434		332
11	306		306

Table 15 Hae III digest profile by *NEB web* cutter
lists the key fragments of the Hae III digest profile for wiltype pSVP2741Aa and the construct pSVP2741Aa_Cry42Aa
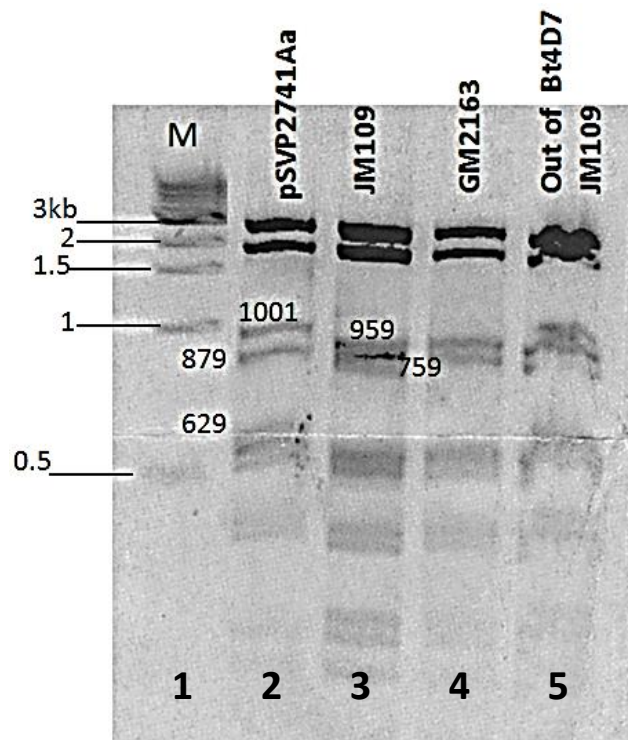
Figure 38 banding profie of construct in agarose gel
Agrose 1.5% gel picture of Hae III digest profile of construct pSVP27*41Aa-42Aa*. All lanes labelled with contents. Lane contained DNA marker Wildtype pSVP2741Aa acts as control in lane 2 1 . Lane 3 contained *JM109 E. coli* strain is where the construct is first confirmed, before it was introduced into *GM*2163 (lane 4) to remove any methylated DNA and later extracted in preparation for transformations with *Bt4D7*. The construct was lastly introduced into *JM109* after mini prep from *Bt* cells to confirm integrity of the construct as shown in lane 5.

Microscopic observations of samples taken from *Bt4D7* expressing the SVP2741Aa_Cry42Aa plasmid confirmed the presence of crystals. The samples were harvested and prepared for characterisation. Crude samples of the hybrid crystals were analysed on SDS PAGE gel. The hybrid expresses both the ~120 KDa (ORF3) and ~80 KDa (ORF2) proteins as seen in the control wildtype sample as shown in figure 39 and was indistinguishable from crude wildtype Cry41Aa.
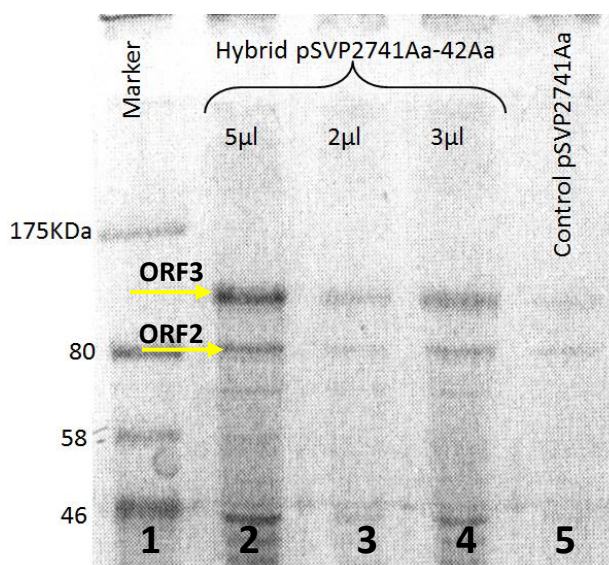
Figure 39 banding profie of construct in agrose gel
SDS-PAGE of crude pSVP2741Aa_Cry42Aa hybrid protein of different volumes 5,2, and 3 µL respectively in lane 2,3, and 4.. Protein marker in lane 1. Control crude Cry41Aa in lane 5. All lanes are labelled with contents and volumes.

The hybrid crystal was solubilised in sodium carbonate pH10.5 and digested with trypsin (1 mg/mL). A soluble form of the hybrid SVP2741Aa_Cry42Aa protein was obtained and evident in the SDS PAGE gel image shown in figure 40. However, despite changing the experimental conditions, it was not possible to obtain a stable protease resistant hybrid protein. The solubilised hybrid is shown in lane 3 where DTT and carbonate have worked together to solubilise the hybrid crystal. A protein thought to be ORF2 and known to have an approximate size of 88 KDa was seen and indicated by a yellow arrow.

Furthermore, an upper band, thought to be ORF3, known to have an approximate size of 120 KDa was also observed and indicated by a yellow arrow. The same band was observed in carbonate and DTT treated control samples of Cry41Aa in lanes 2 and 4.

However, no proteins were observed for hybrid samples incubated with trypsin as indicated by lane 7. Lane 6 was loaded with trypsin treated Cry41Aa. Here a protease resistant protein was observed about ˜80 KDa or slightly less. the shift in size was indicated by a red arrow.
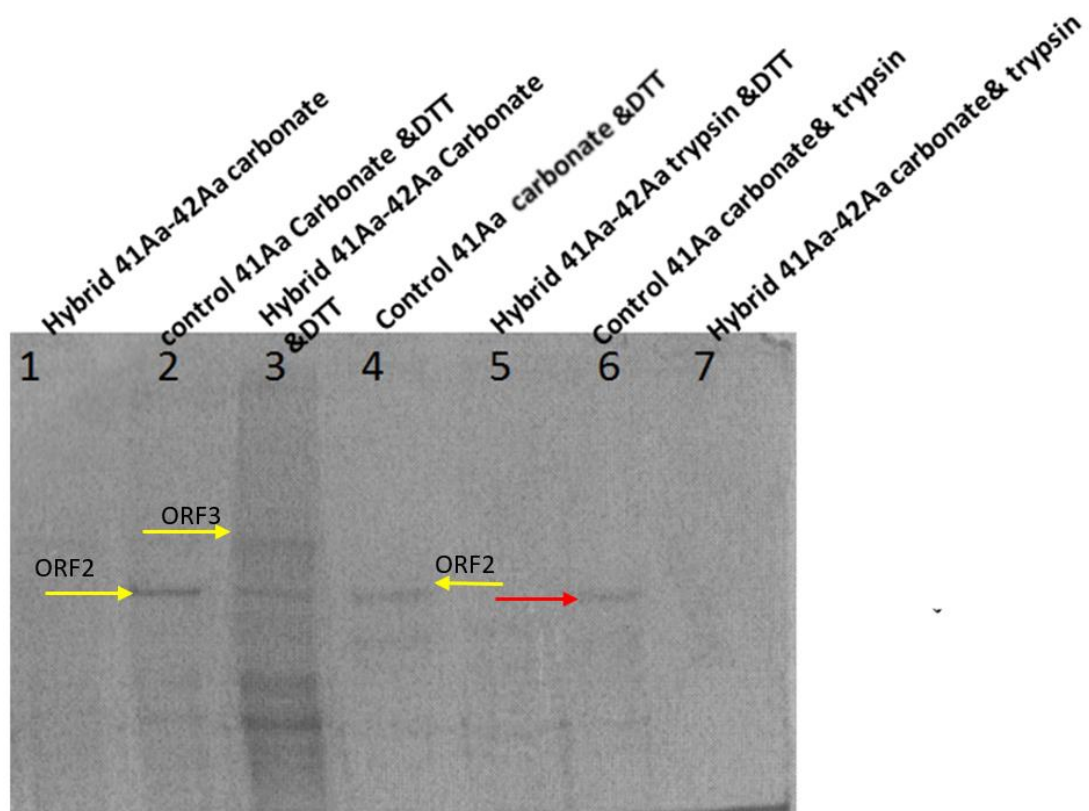
Figure 40 SDS PAGE gel showing pSVP27*41A*a-*42Aa* hybrid protein in various characterisation conditions.
Lane 1was loaded with 5µL of hybrid protein in 50mM carbonate buffer at 10.5pH. Lane 2 was loaded control wildtype Cry41Aa protein in 50mM carbonate buffer at 10.5pH and 0.01mM DTT. Lane 3 was loaded with 5µL hybrid protein in 50mM carbonate buffer at 10.5pH and 0.01mM DTT. Lane 4 was loaded with 5µL of control wildtype Cry41Aa protein in 50mM carbonate buffer at 10.5pH and 0.01mM DTT. Lane 5 was loaded with 5µL of hybrid in 0.1 µg/mL trypsin in 50mM carbonate buffer at 10.5pH and 0.01mM DTT. Lane 6 was loaded with control wildtype Cry41Aa protein in 0.1 µg/mL trypsin in 50mM carbonate buffer at 10.5pH. lane 7was loaded with 5µL of hybrid protein in 0.1 µg/mL trypsin in 50mM carbonate buffer at 10.5pH. yellow arrows indicated ORF2 and ORF3 in control Cry41Aa and solubilised and DTT treated hybrid sample. Red arrow indicated protease resistant core of Cry41Aa

5.3.2 Cry41Aa ORF2 hybrid with Cry1Ie

Cry1Ie is an insecticidal three domain toxin with a protoxin of ~81 KDa. It lacks the C-terminal conserved block 6-8 present in 130 KDa full-length three domain Cry toxins (Song *et al*. 2003). Its specificity has been altered through manipulations that include domain II substitution with *Cry1Aa* resulting in a stable toxin with a new toxicity towards *Spodoptera exigua* (de Maagd *et al.*, 2000).

A pGEM plasmid with the 2160bp ORF gene was created (figure 41). George, 2011 successfully expressed the Cry1Ie protein which was partially solubilised at pH 11. Protease treatment resulted into a ~55 KDa protease resistant core (George, 2011). In an attempt to understand Cry41Aa specificity, a hybrid with Cry1Ie was designed and constructed.

A sequence alignment was carried out and a suitable region for hybrid formation was determined (figure 41). A region of homology was found, primers were designed to amplify conserved block 2, 3, 4, and 5 from Cry1Ie and conserved block 1 from Cry41Aa. As previously mentioned, conserved block 1 and part of conserved block 2 are usually found in domain I which is associated with pore formation. Whilst the remaining part of conserved block 2 and part of conserved block 3 make up domain II which is associated with specificity, binding and receptor recognition. Domain III also associated with binding and pore formation includes part of conserved block 3 and both conserved blocks 4 and 5 (Palma *et al.,* 2014).

In other word this hybrid retained domain I of Cry41Aa, with the intention to investigate it for its ability to exert a toxicity via pore formation. The previous Cry42Aa hybrid did not result in a protease resistant core, perhaps the mutagenesis was too drastic and affected the structural stability of the hybrid. The Cry1Ie hybrid has a large section which comes from Cry1Ie, with the aim that this provides structural stability in the hybrid.

Table 16 below lists the primers used to construct the hybrid. The presence of native Cry41Aa conserved blocks 2, 3, 4 and 5 may ensure better structural stability and thus increase the likelihood that a protease resistant core is obtained post trypsin treatment.

PCR primer list to create construct

Template DNA	Primer name	Oligonucleotide/ Primer sequence	5'-PHO
pBS41Aa	41Aa Hybrid F	5' ATGAATCAAAATTGTAAGTAATAACAATGG3'	Yes
	41Aa Hybrid R1	5'ATGTAGATTCGCAGCTTCTGC3'	Yes
pGEMCry1Ie	1Ie HybridF1	5'ATGAATCAAAATTATAACAACAATGG3'	No
	1Ie HybridR1	5'AAGTTACCTCCATCTCTTTTATTATTAAGATACC3'	No

Table 16 PCR primer list to create construct
 lists PCR primers used to create construct pGEMCry1Ie_Cry41Aa
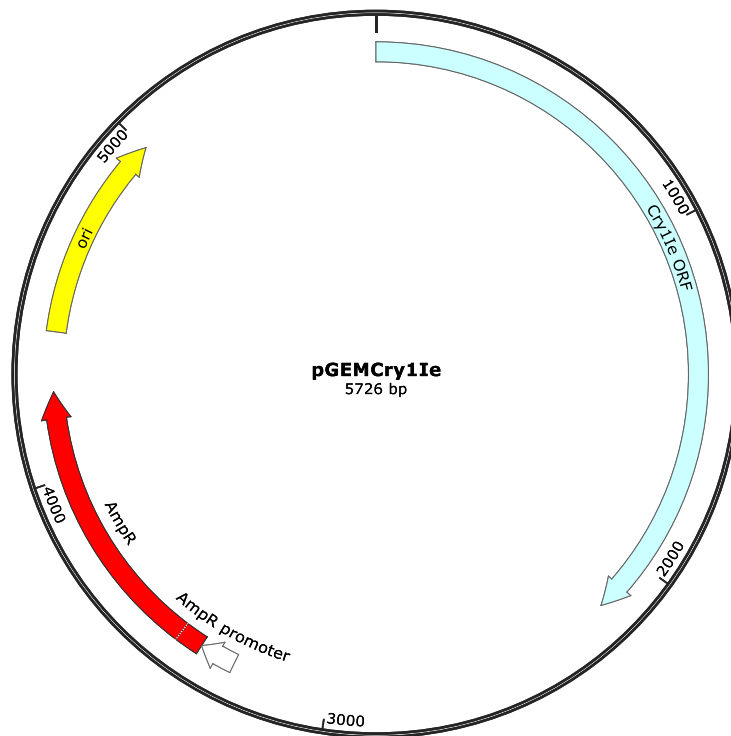
Figure 41 schematic representation of pGemCry1Ie plasmid.
Wildtype pGemCry1Ie plasmid, with Cry*1*Ie ORF in light blue, and an ampicillin resistance gene shown in red. Bacterial origin of replication shown in yellow. Restriction sites not shown for this plasmid

```
CLUSTAL 2.1 multiple sequence alignment

Cry41Aa    MNQNCNNNGYEVLNSGKGYCQPRYPFAQAPGSELQNMGYKEWMNMCTSGDPTVLGEGYSA
cry1Ie     MKLKNPDKHQSLSSNAKVDKIATDSLKNETDIELKNINHEDFLRMSEHESIDPFVSASTI
           *: :  ::  .: ...*    .  .: : .. **:*:.::::.*.   .   : .. :

Cry41Aa    DVRDAVITSINIASYLLSVPFPPAGVAAGILGALLGLLWPTNTQAVWEAFMNTVEALINQ
cry1Ie     QTGIGIAG-----KILGTLGVPFAGQIASLYSFILGELWPK-GKSQWEIFMEHVEELIDQ
           :. .:      . * :: .* ** *.: . :** ***. :: ** **: ** **:*

Cry41Aa    KLDEYARSKAISELNGLKNVLELYQDAADDWNENPGDLRNKNRVLTEFRNVNGHFENSMP
cry1Ie     KISTYARNIALADLKGLGDALAVYHESLESWIKNRNNARATSVVKSQYIALELLFVQKLP
           *:. ***. *::*:** :.* :*::: :.* :* .: * .. * ::: ::  * :.:*

Cry41Aa    SFAVRNFEVNLLPVYAEAANLHLLLLRDAVKFGEGWGMSTDPGAERDDMYRRLRSRTEIY
cry1Ie     SFAVSGEEVPLLPIYAQAANLHLLLLRDASVFGKEWGLSNSQ---ISTFYNRQVERTSDY
           ****  . ** ***:**:*********** **: **:.. . :*.* .**. *

Cry41Aa    TDHCVNTYNQGLQQAKSLQANVSDYSRYPWTQYNQSGGFSYREAKGEYRGTENWNLYNAF
cry1Ie     SDHCVKWYSTGLNNLRGTNA----------------------------ESWVRYNQF
           :****: *. **:: :. :*                            *.*  ** *
```

Extra Loop

```
Cry41Aa    RRDMTILVLDIIAQFPTYDPGLYSRPVKSELTREVYTDIRGTTWRSDANLNTIDAIENRM
cry1Ie     RKDMTLMVLDLIALFPSYDTLVYPIKTTSQLTREVYTDAIGTVHPNASFASTTWYNNNAP
           *:***:**** **:**. :*. ..*:******** **.  . : .*    :*
```

Loop 1

```
Cry41Aa    VGSRQLQLFT---WLTEMKFYIRNTGSITSYTHGDLMVGLEKKIRKTNDNDQWLPLEGQN
cry1Ie     SFSAIESAVVRNPHLLDFLEQVTIYSLLSRWSNTQYMNMWGGHRLEFRTIGGVLNTSTQG
           *    ..    * ::   : .:: ::: : *      :  : .    * . *.
```

Loop 2

```
Cry41Aa    TSYTRIDRPGIELGKNYWYYARTQQWFETRLLQLWANTDVLSLNAGTVGNEFWVRDVPDY
cry1Ie     STNTSINPVTLPFTSRDVYRTESLAGLNLFLTQPVNGVPRVDFHWKFATLPIAS-DNFYY
           :: * *:  : :  ..  * :.:  ::  * *  ..  .:.:  .  :   * *
```

Loop 3

```
Cry41Aa    RNIYARSTRNHFIENHRLSWIKFEP-----VRDNCPFAWPGYKQLSALLFGWTHNSVDLN
cry1Ie     LGYAGVGTQLQDSENELPPETTGQPNYESYSHRLSHIGLISASHVKALVYSWTHRSADRT
           .  . .*: :  **.  .  .:*    :  . :.  . .::**::***.*.*  .

Cry41Aa    NIISQYRITQIPAVKAYWNRGAFSVIRGPGSTGGNLVQLG-TGGEVSVKVRPEQTGSDWY
cry1Ie     NTIEPNSITQIPLVKAFNLSSGAAVVRGPGFTGGDILRRTNTGTFGDIRVNINPPFAQRY
           * *.  ***** ***:   .. :*:**** ***:::: **  .::*. : . :: *

Cry41Aa    RVRIRYAAGSRGRLNVKKYVSSIHASVTYDYNMTMSSSTQGTYNSFQYLDVYNFRLAEPE
cry1Ie     RVRIRYASTTDLQFHTSINGKAINQGNFSATMNRGEDLDYKTFRTVGFTTPFSFSDVQST
           *******: :  :::..   .:*: .         .. *:.:. :   :.*  .:.

Cry41Aa    FEVWLTNESGG-PIWIDKIEFIPLSPIPELPVYPGTYQIVTALNNSSVVTSEEFCMGIGL
cry1Ie     FTIGAWNFSSGNEVYIDRIEFVPVEVT-----YEAEYDFEKAQEKVTALFTSTNPRGLKT
           * :   * *.*  :::**:***:*:.      * . *:: .* :: .:: :.     *:

Cry41Aa    TTRCGVNLWSNNGNTLQKWRFVYNGDQNAFQIKSTPNEDLVLSGSNSGTSVTAETNQNRP
cry1Ie     DVKDYHIDQVSNLVESLSDEFYLDEKRELFEIVKYAKQIHIERNM---------------
            .:       .*    . .*   . .* : .:: *:* . .:: : .

Cry41Aa    NQYWLIEEAGNGYVYLRSKGNPNLVLDVAGTSTANGTNIILWNYNGSTNQKFKLS
cry1Ie     -------------------------------------------------------
```

Figure 41 Protein sequence alignment of Cry41Aa and Cry1Ie.
 Clustal W alignment of the protein sequences of Cry41Aa and Cry1Ie.
conserved blocks of both proteins are underlined and in colour font as follows: conserved block 1 in red, conserved block 2 in orange, conserved block 3 in green, conserved block 4 in pink, conserved block 5 in blue. For Cry41Aa, domain II of Cry41Aa are highlighted as follows: Loop1 in red, loop 2 in yellow, loop 3 in navy and extra loop in pink. The ricin domain is highlighted in teal.

The wildtype pGEM Cry*1Ie* (figure 41) and wildtype BS41Aa (figure 21) plasmids were used as templates to amplify fragments of each Cry toxin to construct pGEMCry1Ie_Cry41Aa plasmid (figure 43).
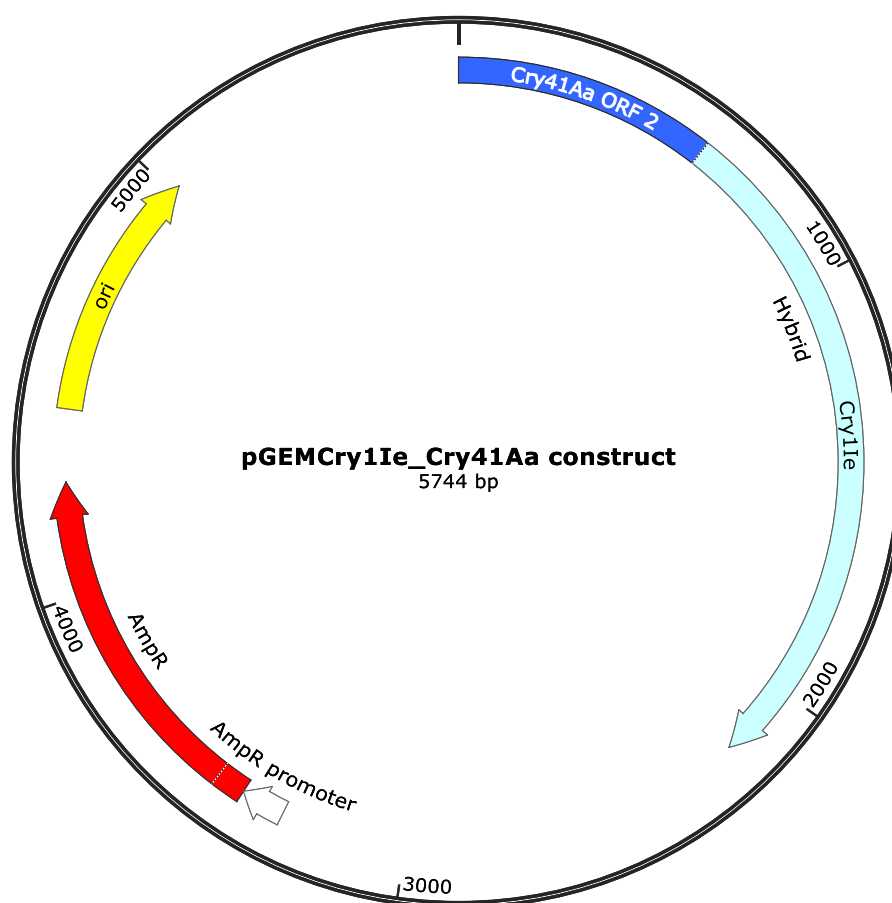
Figure 43 Schematic representation of hybrid plasmid.
Construct plasmid pGemCry1Ie_Cry41Aa with hybrid ORF that is made up of Cry41Aa and Cry1Ie shown in dark and light blue. Ampicillin resistance gene shown in red. Bacterial origin of replication shown in yellow. No Restriction sites shown for this plasmid.

The linear PCR products were run on agarose gel, cut out and gel purified and allowed to ligate overnight. The ligation mixture was introduced into *E. coli GM*2163 and plated onto ampicillin (100 μg/mL) LB agar plates. The plates were incubated overnight at 37˚C.

Six colonies were picked and incubated in ampicillin LB broth for 2-3 h. A Hae III digest was carried out on the recovered plasmids from the six colonies and visualised on agarose gel as shown in figure 44. NEB web cutter software predicted the Hae III digest profiles for both the wildtype and construct as listed in table 17.

Hae III digest profile by *NEB web* cutter

Hae III fragments	Wildtype pGEMCry1Ie (bp)	pGEMCry1Ie_Cry41Aa Construct (bp)
1	2238	2192
2	654	654
3	485	500
4	458	458
5	434	434
6	289	289
7	279	279
8	267	267
9	174	174
10	142	142
11	102	102

Table 17 Hae III digest profile by *NEB web* cutter
lists the key fragements of the Hae III digest profile for wiltype pGEMCry1Ie and the construct pGEMCry1Ie_Cry41Aa

The construct pGEM Cry1Ie_Cry41Aa plasmid had the distinctive 2192bp and 500bp bands which are absent in wildtype pGEMCry1Ie. Figure 39 is the picture of agarose gel analysis of the Hae III digested plasmids extracted from *E. coli* cells. Colony 1 shows distinctive 2192 bp and 500 bp bands.
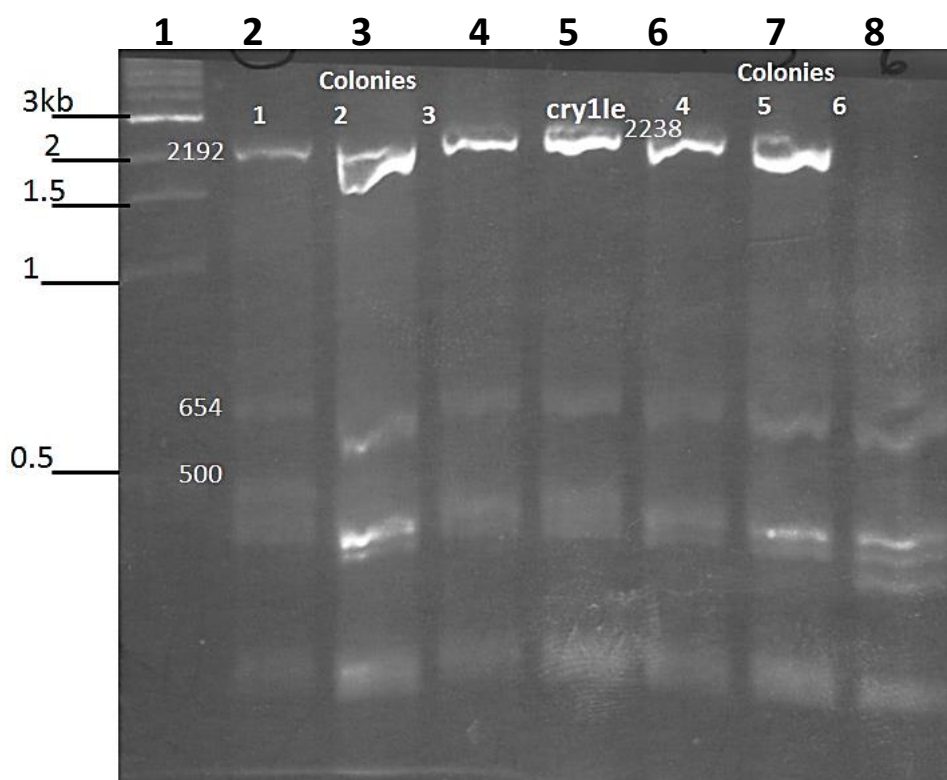
Figure 44 Banding profie of construct in agrose gel
Agarose gel showing Hae III digest profile for plasmids extracted from six bacterial colonies respectively loaded onto lanes 2,3,4,6, 7 and 8. Lane 5 contained wildtype Cry1Ie. All lanes labelled with contents.

The construct was grown in antibiotic LB broth overnight and harvested for the hybrid crystal. The hybrid was subjected to various sodium carbonate and protease conditions in an attempt to obtain a soluble protein. Microscopic observations confirmed the presence of crystal and crude (also referred to as total sample) samples were analysed in SDS-PAGE as shown in figure 45.

Previous Cry1Ie investigations by George *et al*. (2011) indicated that Cry1Ie partially solubilised and resulted in a ~55 KDa core post protease treatment. George *et al*. (2011) carried out the solubilisation and activation of Cry1Ie in one step. In brief, a volume of crude Cry1Ie was centrifuged and superannuant discarded, the pellet was dispersed in

a volume of carbonate, trypsin in carbonate was added and the sample allowed to incubate for 1 h. Contrary to this Cry41Aa is solubilised and activated in two separate steps to get the best yield of activated toxin. Crude Cry41Aa was solubilised in carbonate for 1 h. The sample was centrifuged, and supernatant retained. The solubilised sample was then activated with trypsin and incubated for a further hour.

To address both methods Cry1Ie_41Aa hybrid was processed using both. The one step solubilisation and activation of crude total hybrid samples are shown in lanes 2 to 5. The two-step solubilisation and activation of hybrid Cry1Ie are shown in lane 6 to 10. Crude, samples in lanes 2, 3, and 4 exhibit the major band with a molecular weight of ~81 KDa and are indicated with a yellow arrow. There is no evidence of 55 KDa activated core in trypsin treated crude samples in lanes 4 and 5.

In comparison, lanes 6 was loaded with solubilised only supernatant hybrid and lane 10 was loaded with solubilised and trypsin treated supernatant hybrid, and both did not show any protein bands. Previous observations by George *et al*., (2011) of poor Cry1Ie solubilisation, suggested that this may also be the case for the hybrid. Additionally, it is likely that the hybrid was not stable in trypsin. Despite altering the carbonate pH and trypsin concentration, a solubilised hybrid capable of resulting in a trypsin resistant core was not obtained.
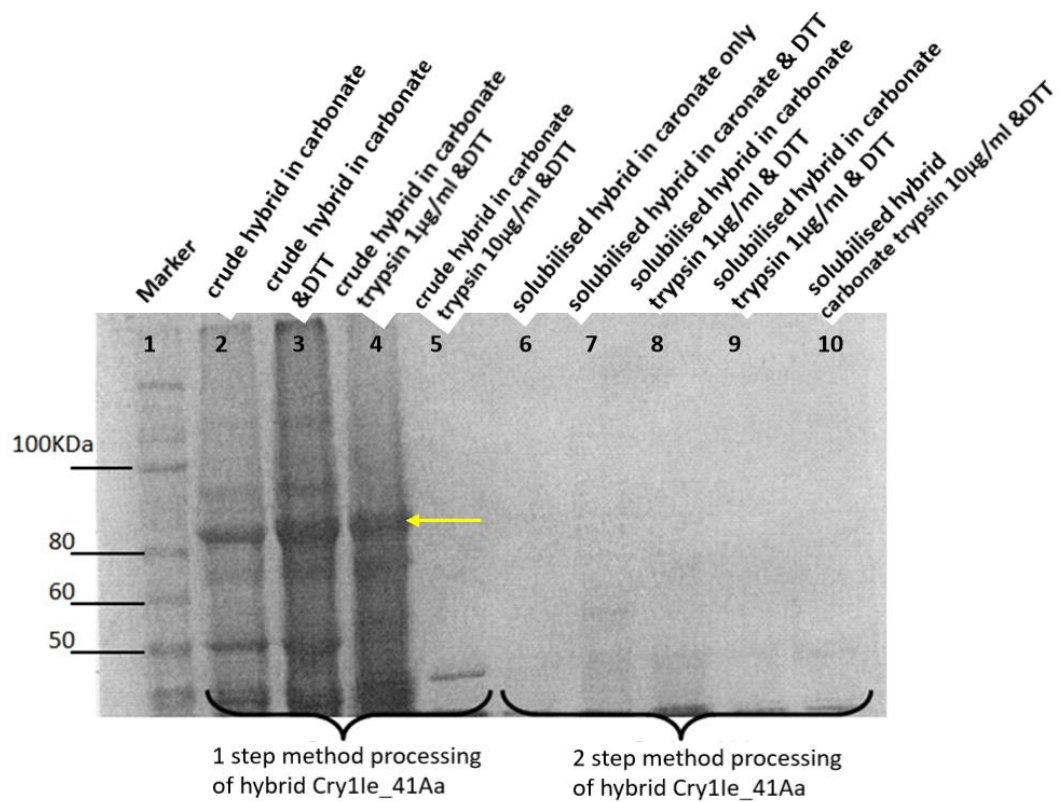
Figure 45 SDS-PAGE gel of hybrid pGEMCry1Ie_Cry41Aa.
All lanes labelled with content. Lane 1 contained protein marker. Lane2 was loaded with 5µL of crude hybrid in 50mM carbonate buffer at 11pH. Lane 3 was loaded with 5µL crude hybrid in 50mM carbonate buffer at 10.5pH and 0.01mM DTT. Lane 4 was loaded with 5µL of crude hybrid in 0.1 µg/mL trypsin in 50mM carbonate buffer at 11pH and 0.01mM DTT. Lane 5 was loaded with 5µL of crude hybrid in 1 µg/mL trypsin in 50mM carbonate buffer at 11pH and 0.01mM DTT. Lane 6 was loaded with 5µL of supernatant hybrid in 50mM carbonate buffer at 11pH. Lane 7was loaded with 5µL supernatant hybrid in 50mM carbonate buffer at 10.5pH and 0.01mM DTT. Lane 8 was loaded with 5µL of supernatant hybrid in 0.1 µg/mL trypsin in 50mM carbonate buffer at 11pH and 0.01mM DTT. Lane 9 was loaded with 5µL of supernatant hybrid in 0.1 µg/mL trypsin in 50mM carbonate buffer at 11pH and 0.01mM DTT. Lane 10 was loaded with 5µL of supernatant hybrid in 10 µg/mL trypsin in 50mM carbonate buffer at 11pH and 0.01mM DTT.

5.4 Discussion

The application and creation of Cry hybrids has been used to enhance toxicity, create new specificity and elucidate modes of action (Abdullah and Dean, 2004; Abdullah *et al*., 2003; Pardo-López *et al*., 2009; Dean. H and Sylvis L, 2006; Dean *et al*., 1996). Currently there is more than one model to explain the mode of action of Cry toxins (Vachon *et al*., 2012, Pardo-Lopez *et al*., 2013) However, all theories on the mode of action agree that specific binding must take place before toxicity can commence (Likitvivatanavong *et al*., 2009).

Mutagenesis has been applied to investigate the role that specific regions of a given protein play in its mode of action. For insecticidal 3-domain toxins, the assembly of hybrids has mainly centred on domain swapping. This led to the creation of novel toxins with a wider target spectrum or a significant increase in toxicity to the target cell compared to that of the wildtype (Pardo-López *et al*., 2009).

Initially the research plan was to carry out a Cry41Aa loop 3 exchange that would result in the removal of the wildtype loop 3 and the introduction of the insecticidal loop 3. It was thought that this exchange may result firstly in a loss of specificity towards HepG2 by Cry41Aa and or secondly a gain of specificity towards another insect cell. Abdulla *et al*. (2004) employed a similar experimental design when loops 1 and 2 of Cry4Ba (toxic to *Anopheles* and *Aedes*) were deleted and replaced with loops 1 and 2 of Cry4Aa (toxic to *Culex*). Despite the production of proteotically stable hybrids proteins, they did not

exhibit any toxicity towards susceptible insects (Abdullah and Dean, 2004). Mutagenesis of Cry4Ba involved the introduction of loop 3 of Cry4Aa; this resulted in a hybrid that not only gained toxicity to *Culex* but exhibited greater toxicity than wildtype Cry4Ba to *Ades ageypti*. It has been suggested that the natural evolutionary mechanism can act on Cry toxins and subsequently generate new specificities (Bravo, 1997; de Maagd *et al*., 2001; Wu *et al*., 2007).

The synthetic creation of hybrids via mutagenesis can be seen as mirroring natural selection in Cry proteins (Caramori *et al*., 1991). For example, Cry1Ab, Cry1Ac, Cry1Ba and Cry1Ea have little or no toxicity towards *Spodopetera exigua*, but they gained toxicity after swapping native domain III with domain III of Cry1Ca which targets this insect. Furthermore, the Cry1Ab hybrid has a significantly higher toxicity towards *S. exigua* compared to Cry1Ca (de Maagd *et al*., 2000).

In another study by Abdullah *et al*. (2003) the loops of domain II Cry1Aa (toxic to lepidopteran) were designed to resemble mutants of Cry4Ba and the resulting Cry1Aa hybrid gained toxicity towards *Culex pipiens* larvae. Such findings suggest that the domain II loop can alter the specificity of a given toxin (Dean and Sylvis ., 2006).

Unfortunately, attempts to make hybrids with Cry41Aa did not result in the generation of trypsin resistant protein cores, the hybrids were often unable to solubilise well and or were not stable in trypsin. Previous studies have highlighted the possibility that hybrids are susceptible to additional cleavage by host bacterial proteases (Abdullah and

Dean, 2004). This may explain the sudden absence of the hybrid protein when analysed in SDS PAGE gels. This finding suggests that Cry41Aa is sensitive to the slightest of structural changes. It seems that any subtle changes on the DNA of Cry41Aa appear to translate into significant structural changes.

It is not known if Cry41Aa is toxic to other vertebrate or insecticidal cell lines. Without a protease resistant protein, it is difficult to shed further light on Cry41Aa specificity. At this stage. The findings suggest that mutagenesis on Cry41Aa by loop exchange or creation of hybrids was too drastic change and its structure too compromised. In a study carried out to investigate the effects of loop exchange between native domain II loops of Cry4Aa, it was revealed that exchange between loop 1 and 3 decreased toxicity to *Culex* but did not abolish toxicity in bioassay trails. However, a loop 2 exchanges eradicated toxicity completely and led to the proposal that it is loop 2 of Cry4Aa that is essential for mosquitocidal activity (Howlader *et al.*, 2009).

A study by Masson *et al.* (1994) proposed that N terminal regions do not affect specificity, after it was observed that N terminal exchange between Cry1Ac and Cry1Ea did not introduced novel specificity. The Cry41Aa_Cry1Ie hybrid did introduce an N terminal region of Cry1Ie to Cry41Aa. Unfortunately, the resulting hybrid was unstable in both sodium carbonate and protease conditions. Indeed, the only successful hybrid made was the solubilised trypsin treated pGEM1Acloop3 hybrid. Indeed, the only successful hybrid made was the solubilised trypsin treated pGEM1Acloop3 hybrid. This hybrid was made up of insecticidal Cry1Ac with loop 3 of Cry41Aa. Once solubilised and

activated it resulted in a protein resistant core of ~ 65 KDa. Preliminary cell assay analysis was carried out on HepG2 cells with pGEM1Acloop3 hybrid to determine if it gained cytocidal activity. Observations indicated that the hybrid was not toxic to HepG2 cells. However, a study conducted both *in vivo* and *in vitro* cultured insect cell lines and dissociated midgut epithelial cells has concluded that hybrids can exhibit different toxicity which is dependent on bioassay systems in place (Masson *et al*., 1994). To shed light on the specificity of Cry41Aa a different mutagenesis approach was required in order to obtain a stable recombinant protein.

6.0. Production, purification, and characterisation of Cry41Aa loop mutants

6.1 Introduction

Amino acid substitutions were used to probe the mechanism of action of Cry insecticidal toxins (Dean *et al*., 1996). The theory that domain I is attracted to negatively charged membrane surfaces was researched in the early nineties. The surface residues of exposed loops of Cry1Ab in domain I were substituted and the study concluded that a positively charged residue such as arginine or a neutral residue such as alanine had a stronger membrane insertion than the negatively charged aspartic acid wildtype residue (Dean *et al*., 1996).

Alanine substitutions were made in domain II loops, in a study by Wu *et al.* (1996) which explored the toxicity of Cry3Aa and its ability to bind to receptor gut epithelia receptors of *Tenebrio molitor* beetle larvae. Alanine substitutions in loop I and loop 3 of domain II indicated a reduced binding affinity and toxin stability. However, a block of alanine substitution made in loop 3 despite having a lower binding affinity resulted in a higher toxicity compared to wildtype Cry3Aa. This was noted to be due to the mutants ability to have an increased membrane insertion compared to that of the wildtype (Wu and Dean, 1996).

A study by Garcia-Robles *et al.* (2012) explored the specificity and binding of insecticidal Cry3Aa towards the Colorado potato beetle. Site directed mutagenesis was carried out in loop 1 of Cry3Aa. The mutagenesis carried out corresponds to a previously identified ADAM metalloprotease recognition motif thought to take part in membrane associated proteolysis and toxicity in the Colorado potato beetle. The study concluded that interactions with ADAM through the Cry3Aa recognition motif led to Cry3Aa proteolysis, and that this is vital for Cry3Aa toxic action in target insect.

Further studies were carried out in nCry3Aa that investigated the receptor recognition and binding of Cry3Aa by *T. molitor*. One study introduced an alanine substitution in loop1 and 2 where a loss of toxicity was observed potentially due to reduced receptor binding. However, a loop 3 block alanine substitution resulted in an increase in toxicity due irreversible binding where the toxin is thought to insert itself in the membrane of target cells (Wu and Dean, 1996, Nachimuthu and Polumetla Ananda, 2004).

Rajamohan *et al.* (1996) explored the effects of alanine substitutions on loop 2 and 3 of Cry1Ab, which resulted in the loss of toxicity to *Manduca sexta* and *H. virescens* as consequence of reduced binding affinity to BBMV (Rajamohan *et al*., 1996a; Rajamohan *et al*., 1996c; Nachimuthu and Polumetla Ananda, 2004). This was later supported by findings that suggested that loops α8 and 2 of Cry1Ab interact with *M. Sexta* Bt-R1 receptor. (Gomez *et al*., 2003).

In this study, domain II loops 1, 3, and the extra loop from Cry41Aa underwent residue substitutions in the quest to understand its specificity towards mammalian HepG2 cell line. This mutagenesis approach involved the substitution of specific residues (i.e. amino acid substitutions) in domain II loops of Cry41Aa. The residue substitutions on domain II loops of Cry41Aa took place in residues of the loop sequences as identified in chapter 4.

Substitutions on the three loops of Cry41Aa were carried out in concurrence with one another so that each mutagenesis could provide information that would go on to build a picture of which residues and possibly how Cry41Aa is able to target HepG2 cell.

Bioinformatics has predicted with confidence the sequences that make up loop1 and loop 3 of Cry41Aa. It had highlighted the uniqueness of the extra loop in category of 3-domain Cry toxins and predicted it to have a secondary structure that is mostly made up of strands. Previous attempts to delete all or some of the extra loop have resulted in unstable recombinant proteins (Krishnan *et al*., 2017).

Studies that explore insecticidal 3-domain Cry toxins have typically investigated the exposed loops of domain II. The literature has highlighted the importance of these loops in the toxin's ability to bind to and interact with target cells. The sequences of loop1 and loop 3 of Cry41Aa have been predicted as follows.

Loop1 is a short four-residue sequence: ^{384}SITS$^{387.}$ As it is the shortest loop being investigated it allows for several residue substitutions in attempt to collate as much

information possible on the specificity of Cry41Aa. Loop 3 of insecticidal 3-domain Cry proteins has been directly implicated in the mode of action of 3-domain toxins (Pacheco *et al*., 2009). Loop 3 of Cry41Aa is predicted to be 15 residues long sequence: [503]VRDNCPFAWPGYKQL[517]. Previous attempts to exploit loop 3 to create hybrids did not result in protease resistant protein (chapter 5.2), substitutions here can aid to narrow down sequences or residues that play a key role in Cry41Aa ability to target HepG2 cells. The three loops selected for investigation are shown in figure 46. The side chains of the residues chosen for substitution are shown for each loop of domain II in Cry41Aa.

Since little is known about Cry41Aa it is difficult to hypothesise the effect of mutagenesis on toxicity which may improve, reduce or get completely knocked out.

This section details how domain II loop constructs were designed and made. These were made using the general steps as outlined in introductory figure 20, chapter 5.1. All constructs have a PCR design that is specific to the mutation made. A small plasmid such as the wildtype pBS41Aa was commonly used as a DNA template for the PCR reaction. The wildtype plasmid contains ORF2 of Cry41Aa that is flanked by restriction sites (BamHI and XhoI) as well as an ampicillin resistance gene (figure 21).

The PCR design is shown and details each construct's design. From then on, all construct proceeded through similar lab procedures and check points. All linear PCR products are excised from agarose gels, purified and allowed to ligate overnight. The ligation mixture was introduced into an *E. coli* strain (usually *JM109).* The transformation mixture was poured onto ampicillin LB agar plates for overnight incubation at 37 $^{\circ}$C. Surviving

colonies were picked and grown either in antibiotic LB for 2-3 h or by re-streaking onto antibiotic LB agar plates for overnight incubation. Mini preps were carried out to extract potential constructs from bacterial cells. These underwent Hae III digests and gel images of the banding profile were compared to predicted NEB web software Hae III profiles.

 Once the constructs were confirmed by Hae III digest profiles and sequencing, the mutant ORF2 was subcloned into a *Bt* expression vector. This step involved a double digest by BamH and XhoI enzymes on both the construct and the *Bt* shuttle vector (pSVP2741Aa figure 28). The fragments were allowed to ligate overnight to form the final *Bt* expression construct. A series of mini preps, Hae III digests and sequencing took place to confirm the constructs and their integrity, as they undergo a number of various strains *E. coli* and *Bt4D7* bacterial transformations.

This section shows in detail one or two constructs as examples of the processes and checkpoints that all Cry41Aa domain II loops constructs were subject to it also shows evidence of the recombinant proteins produced as a result of the mutagenesis carried out in loops of Cry41Aa, and their preparation for cell assays. The chapter begins with mutagenesis carried out on loop 1 of Cry41Aa*.* Followed by mutagenesis in the extra loop of Cry41Aa, and finally in loop 3 of Cry41Aa.
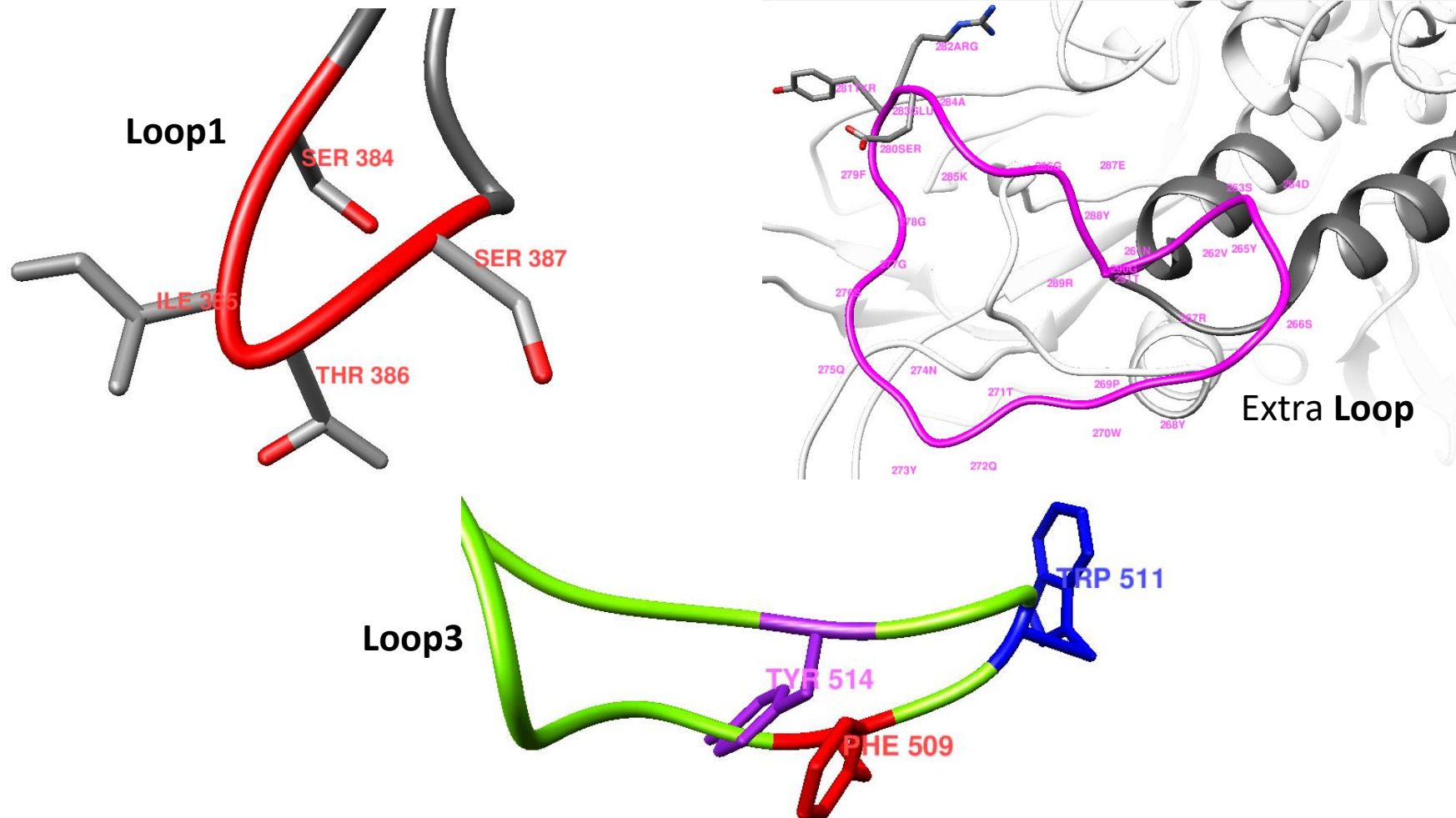
Figure 46 UCF chimeria visualised putative loops of Cry41Aa **on which mutagenesis was carried out.**
 Loop1 shown in red and has the sequence [384]SITS[387] , all four residues underwent substitutions. Loop 3 shown in green and has the sequence [503]VRDNCPFAWPGYKQL[517]   the three hydrophobic residues which were substituted are shown in loop 3 F509 (red), W511 (blue), and Y514 (purple). extra loop is shown in pink and has the sequence [261]NVSDYSRYPWTQYNQSGGFSYREAKGEYRGT[291]. Substitutions in [280]SYRE[283] were carried out on the extra loop.

6.2.1 Establishing optimum solubilisation and activation conditions

The optimum solubilisation and activation conditions for wildtype Cry41Aa were investigated and later set as testing standard for all loop recombinants. *Bt* cells were harvested, crystals were analysed and visualised on SDS PAGE gel. In this study, bioinformatic analysis of Cry41Aa has revealed that it has a similar structure to insecticidal 3-domain Cry toxins known to have separate ORF or more than one ORF for crystallisation. These split genes encode products that resemble the C-terminal half and the N-terminal half of 135 KDa or larger than 60 KDa toxins (Adalat *et al*., 2017, Bietlot *et al*., 1990). Cry41Aa is known to have ORF2 and ORF3 thought to be responsible for an ~80 KDa and ~120 KDa protoxin respectively. Yamashita *et al*. (2005) has argued that both ORF2 and ORF3 are necessary to encode and crystallise the cytotoxic crystal and that previous attempts to produce Cry41Aa crystals without ORF3 did not result in crystal formation (Yamashita *et al*., 2005, Krishnan, 2013).

All 3-domain Cry toxins have distinct globular domains, domain I, II and III are known to produce protoxins of ~65 or ~135 KDa where the highly conserved C-terminal aids to crystallise the protoxin by forming intermolecular disulfide bonds and non-covalent interactions which stabilise the crystal structure (Adalat *et al*., 2017). Once ingested the alkaline insect gut provides the reducing environment to dissociate the cross-links bonds effectively solubilising the crystal. The soluble protoxin is subsequently processed by insect gut proteases (Bravo *et al*., 2007). An *in vitro* setup was established to mimic the pH and protease conditions present in an insect gut in order to solubilise and activate

Cry41Aa and its mutants. Previous studies have observed proteolysis of endotoxins as a result of native bacterial proteasese released during the in *vitro* solubisation stage , the Cry41Aa crystals and its mutants were thus washed with 1M sodium chloride to remove absorbed bactrial proteases (Zalunin *et al.*, 2004).

The harvested Cry41Aa crystals and its crystal mutants were solubilised in sodium carbonate pH 10.5. The reducing agent 2-mercaptoethanol and dithiothreitol (DTT) was also added to disrupt the disulfide bonds. Crude total sample of Cry41Aa crystals were exposed to different carbonate and trypsin conditions and incubated for 1h in a 37˚C water bath.

The samples were then analysed on an SDS PAGE gel as shown in figure 47. Initially both solubilisation and activation were carried out in a single 1 h incubation step. The gel indicates that there was no significant difference between the Cry41Aa crude control (crystals) sample and the Cry41Aa crystals that have been incubated with 50mM sodium carbonate pH 11, trypsin (2 mg/mL), and DTT (2.5mM). Both protoxins ORF3 (~120 KDa protein) and ORF2 (~80-88 KDa protein) are present in these conditions and indicated by yellow arrows in the gel. It indicated that on this instant Cry41Aa crystals did not solubilise well and activation was also inadequate.

Yamashita *et al,* (2005) proposed that the proteinase K activated 64 KDa toxic protein is derived from the 88 KDa protoxin, and it is this 64 KDa peptide which is responsible for the cytocidal activity to HepG2 cell lines. The gel did not clearly show the activated ˜60 KDa toxin.

The characterisation protocol of Cry41Aa was adjusted to improve solubilisation and activation of the toxin. These were carried out in separate steps independent of one another in order to obtain first the solubilised protoxin and then to activate it with trypsin. This is the 2-step method. A number of conditions were tested. The pH levels, DTT and trypsin concentration were optimised. Figure 48 is the SDS PAGE gel picture of solubilised activated Cry41Aa in various trypsin concentrations.

Crude Cry41Aa was incubated for an hour in 50 mM sodium carbonate pH 10.5 with DTT in a water bath at 37˚C. The sample was then centrifuged, and the supernatant was retained. The solubilised Cry41Aa was activated by different concentrations of trypsin to establish the optimum concentration of trypsin for activation. The 1 mg/mL trypsin activated sample gave a clear ˜80 KDa or less major protein and an ˜60 KDa minor protein both indicated by yellow arrows in gel. These two proteins are thought to originate from the solubilised and activated ORF2 encoded crystal. The 120 KDa protein (ORF3) is no longer detected in the trypsin activated gel lanes but is clearly noted in the carbonate-DTT solubilised sample prior to trypsin activation. Background proteins were

noticeably reduced compared to single step solubilisation activation gel picture (figure 47).

Therefore, all crystals including wildtype were solubilised and activated in a 2-step process that took 2 h. First the crude samples were incubated for an hour in 50mM sodium carbonate pH 10.5 and DTT (2.5mM) in a 37 ^{o}C water bath. The samples were centrifuged, and supernatant retained. Followed by the activation step followed where trypsin (1 mg/mL) was added and the samples were incubated for 1 h in a 37 ^{o}C water bath. This procedure was applied as standard to all mutant crystals as well as wildtype Cry41Aa crystal. The activated form of the toxins was ready for further analysis.
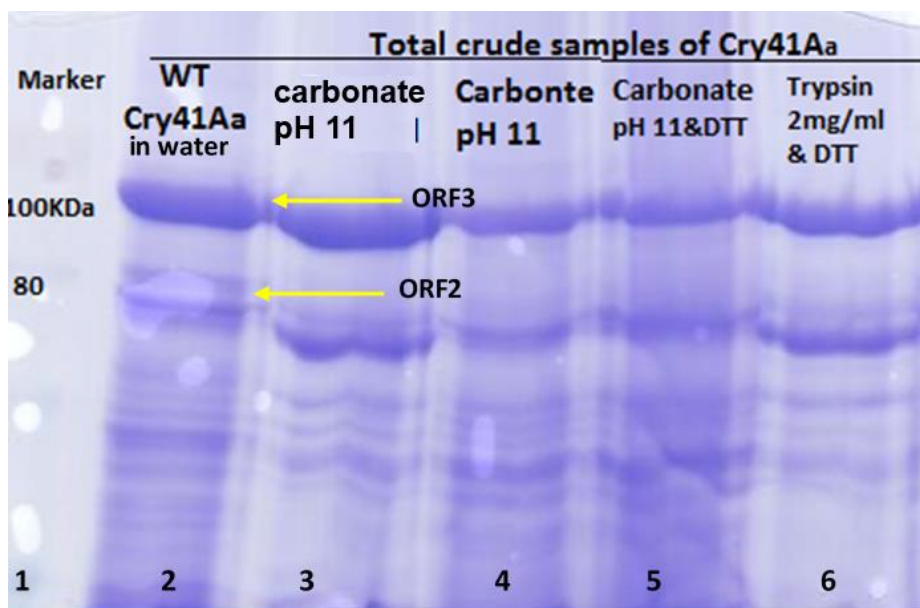
Figure 47 7.5% SDS PAGE gel of crude total (crystal) Cry41Aa

Cry41Aa was incubated for I h in a 37°C water bath in varying different pH, trypsin and DTT conditions. Lane 1 contained protein marker. Lane 2 contained wildtype crude Cry41Aa in water. Lane 3 contained crude Cry41Aa treated with 2 mg/mL trypsin in carbonate. Lane 4 contained crude Cry41Aa incubated with carbonate pH 11. Lane 5 contained crude Cry41Aa incubated with carbonate pH 11 and DTT. lane 6 contained crude Cry41Aa incubated with trypsin 2 mg/mL in carbonate and DTT. All lanes labelled with content.
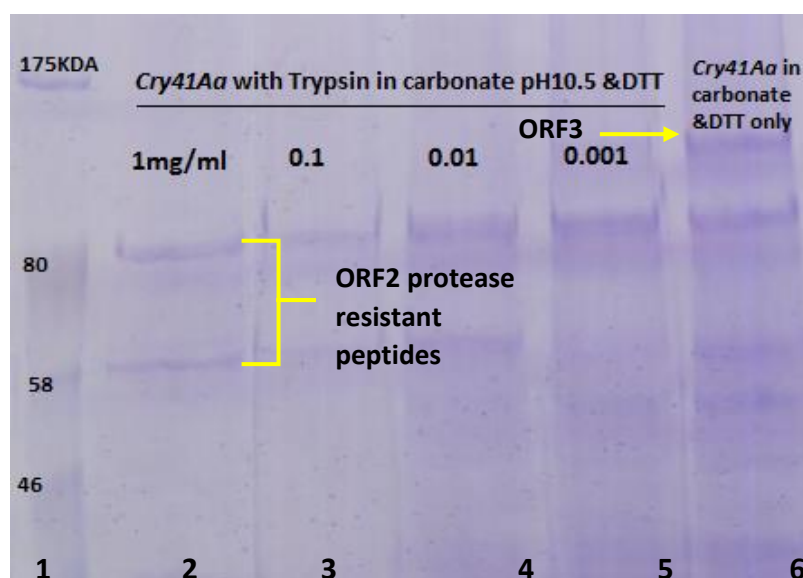
Figure 48 7.5% SDS PAGE gel of a solubilised Cry41Aa supernatant treated with different concentrations of trypsin in carbonate pH 10.5 and DTT (2.5mM).
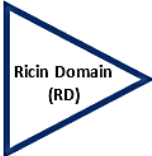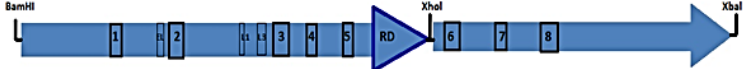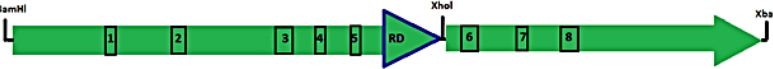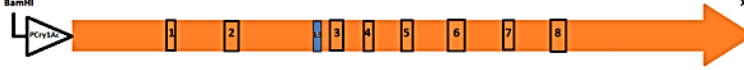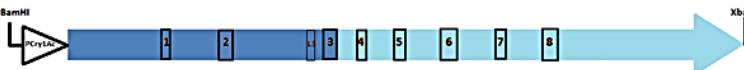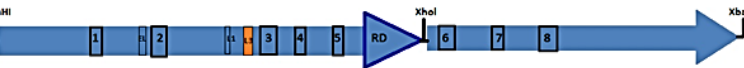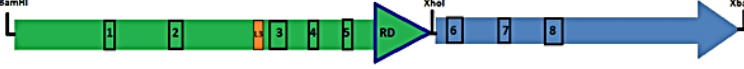
lane 1 contained protein marker. Cry41Aa was activated by different trypsin in carbonate concentrations, 1 mg/mL (lane 2), 0.1 mg/mL (lane3, 0.01 (lane 4, 0.001 (lane5). Lane 6 contained solubilised Cry41Aa in carbonate pH 10.5 and DTT.

6.2.2 Assessment of Cry41Aa mutants and their effect on cell lines

In an effort to investigate the specificity and toxicity of Cry41Aa towards HepG2 cells, a number of recombinants mutants from loop1, 3 and extra loop were created, and their cytotoxicity to a number of different human cell lines were investigated.

Cell assay analysis carried out on HepG2 cells incubated with Cry41Aa and its recombinant toxins to establish the cytotoxicity of the activated proteins. Cell viability assay that measure metabolic activity of viable cells was applied to establish toxicity. Membrane permeability assay was applied to confirm nuclear membrane breakdown. Western blots analysis was applied to depicts the activation of p38 MAP kinase in toxin treated cells. This particular pathway is activated as a consequence of membrane damage by pores (Zhang *et al.*, 2005, Ratner *et al.*, 2006). Finally, the presence and type of potential pores as a consequence of the effect of Cry41Aa and its loop recombinant was also addressed. All Cry41Aa recombinant toxins were characterised against Cry41Aa, these recombinants formed crystals which were harvested, solubilised and activated with trypsin in the same manner as Cry41Aa. Table 18 below summarises the constructs made to investigate the specificity of Cry41Aa to HepG2 cells.

Schematic table of Cry toxin mutations

Legend: Cry41Aa | Cry41Ab | Cry42Aa | Cry1Ac | Cry1Ie | Cry41 loop mutation | Ricin Domain (RD)

Construct diagram	Construct Name	ORF2	ORF 3	Progression/ characterisation check points					
				Host cell	Mutation made	Express	Solubilise	Trypsin treated	Cell assay
	Svp2741Aa XhoI	41Aa	41Aa	E. coli &BT	None-wild type	✔	✔	✔	Toxic to HepG2
	Svp2741Ab XhoI	41Ab	41Ab	E. coli &BT	None -wild type	✔	✔	✔	Toxic to HepG2
	svp27he41Aa Nhe	42Aa	41Aa	E. coli &BT	None -wild type	✔	Degrades	–	–
	pGEM1Ac 41Aa loop3	Cry1Ac		E. coli	Substitute WT loop with loop 3 of Cry41Aa	✔	✔	Not tested on susceptible insect	Nontoxic to HepG2.
	pGEMCry1Ie 41Aa hybrid	Cry41Aa_Cry42 hybrid		E. coli		✔	✔	✔	Degrades unstable	–
	Svp2741Aa Cry1Ac loop 3	41Aa	41Aa	E. coli & Bt	Substitute WT loop with loop 3 of Cry1Ac	✔	✔	Degrades unstable	–
	Svp2741AbCry1Ac loop 3	41Ab	41Aa	E. coli & Bt	Substitute WT loop with loop 3 of Cry1Ac	✔	✔	Degrades unstable	–

Construct								
42Aa_41Aa hybrid	42Aa & 41Aa	41Aa	*E. coli & Bt*	936 bp=42Aa 1524bp=41Aa	✔	✔	Degrades unstable	–
Extra loop alanine substitution	41Aa	41Aa	*E. coli &BT*	SYRE substituted to AAAA	✔	✔	✔	Not toxic
Loop1 degenerate residues substitution	41Aa	41Aa	*E. coli &BT*	SITS substituted to CVSC CLAC GLAC	✔	✔	✔	Toxic to HepG2
Loop3 F509A substitution	41Aa	41Aa	*E. coli & BT*	Alanine substitution	✔	✔	✔	Non toxic to HepG2.
Loop3 F509L substitution	41Aa	41Aa	*E. coli &BT*	Alanine substitution	✔	✔	✔	Non toxic to HepG2.
Loop3 F509Y substitution	41Aa	41Aa	*E. coli &BT*	Tyrosine substitution	✔	✔	✔	Toxic to HepG2
Loop3F509S substitution	41Aa	41Aa	*E. coli &BT*	Serine substitution	✔	✔	✔	Non toxic to HepG2.
Loop3 F509C substitution	41Aa	41Aa	*E. coli &BT*	Cystine substitution	-Weak expression -absent ORF2 in all samples	–	–	–
Loop3 F509W substitution	41Aa	41Aa	*E. coli &BT*	Tryptophan substitution	✔	✔	✔	Non toxic to HepG2.

Loop3 F509Y substitution	41Aa	41Aa	*E. coli* &BT	Tyrosine substitution	✔	✔	Degrades unstable	-
Loop3 F509F substitution	41Aa	41Aa	*E. coli* &BT	Phenyl-alanine substitution	✔	✔	✔	toxic to HepG2.
Loop3 W511A substitution	41Aa	41Aa	*E. coli* &BT	Alanine substitution	✔	✔	✔	Non toxic to HepG2.
Loop3 W511F substitution	41Aa	41Aa	*E. coli* &BT	Phenyl-alanine substitution	✔	✔	✔	toxic to HepG2.
Loop3 W511Y substitution	41Aa	41Aa	*E. coli*	Tyrosine substitution	✔	✔	Degrades unstable	-
Loop3 Y514A substitution	41Aa	41Aa	*E. coli* &BT	Alanine substitution	✔	✔	✔	toxic to HepG2.

Table 18 A summary of the constructs made to invetigate the specificty of Cry41Aa toxin.
List of Contructs name along with details of ORFs, host bacterial cell, mutagensis, and whether they expressed proteins. The progression of expressed proteins from each constrcut through the stages of solublisation, trypsin activatiation, and toxicity on HepG2 cell lines.

6.3 Residue substitutions in loop 1 of Cry41Aa.

Bioinformatics was used to predict loop1 of Cry41Aa. The native sequence [384]SITS[387] underwent site directed mutagenesis with the aim of creating various mutants in an effort to understand Cry41Aa specificity towards HepG2 cells.

 Four degenerate nucleotides were included in the forward primer to make three possible substitutions, making the number of possible mutant outcomes $3^4=81$ as detailed in table 19. The pBS41Aa plasmid acted as template DNA. Once PCR products were purified and allowed to self-ligate, *E. coli* transformations were carried out.

Table 20 lists the PCR primers for loop1 substitutions. Colonies with potential degenerate constructs were first identified by RSS. Bacterial colonies were picked and lysed 2% NaOH solution. The intact plasmids were separated on an agarose gel, large plasmids with insert travel slower than smaller plasmids no insert control (Law and Crickmore, 1997). In this instant loop1 mutant plasmids were the same size as wildtype BS41Aa and appeared to run equally when loaded in an agarose gel. Constructs that appeared similar to BS41Aa on RSS agarose gel were incubated overnight, centrifuged, pellet resuspended and mini prepped for further analysis. The extracted plasmids were subject to Hae III digests and were indistinguishable from wildtype Cry41Aa Hae III digest profile on an agarose gel. Constructs were confirmed by sequencing.

The mutant plasmids were introduced into *Bt* where the recombinant proteins were expressed. The resulting loop1 mutants' proteins were CVSC, CLAC, and GLAC, their toxicity towards HepG2 was investigated.

Native residues OF CRY41AA LOOP1

	5' T G S I T S Y T 3'
Codons	5' ACG GGT AGC ATT ACT AGC TAT ACT 3'
Degenerate codons	BGC BTT BCT BGC
Outcomes	[384] Serine: AGC to BGC: Glycine, Cysteine, Arginine [385]Isoleucine ATT to BTT: Phenylalanine, Valine, Leucine [386]Threonine ACT to BCT: Alanine, Serine, Proline [387]Serine AGC to BGC: Glycine, Cysteine, Arginine

Table 19 PCR primer list to create constructs
primer design for the substitutions for each residue in loop1 of Cry41Aa following the introduction of degenerate codons. Native residues and codons are highlighted in red.

primer list to create constructs

Template DNA	Primer name	Oligonucleotide/ Primer sequence	5'-PHO
pBS41Aa	JTLoop1F	5' GGTBGCBTTBCTBGCTATACTCACGGTC 3'	Yes
	JTLoop1R	5' GGTATTCCTTATATAAAATTTCATTTCTG 3'	Yes

Table 20 PCR primer list to create constructs
lists PCR primers used to create loop1 substitution construct

6.3.1. Establishment of the cytotoxicity of loop1 Cry41Aa recombinants

The predicted loop1 of Cry41Aa is a relatively short surface loop with the sequence [384]SITS[387]. Substitutes were made using degenerate nucleotides and successful

recombinants were recovered. These recombinants were characterised against the wildtype Cry41Aa. Recombinants formed crystals that were harvested, solubilised and activated with trypsin in the same manner as Cry41Aa. The recombinants behaved similar to Cry41Aa and were indistinguishable on an SDS PAGE gel. Initial cell assay data indicated that all loop1 substitutes retained toxicity towards HepG2 cells. Since the mutant protein's concentrations were not standardised, it was not possible to establish whether some mutants were significantly more or less toxic than Cry41Aa. It was clear that the loop1 substitutions did not knock out toxicity suggesting that this putative loop may not play a key role in the specificity or toxicity of Cry41Aa as shown in figure 49.
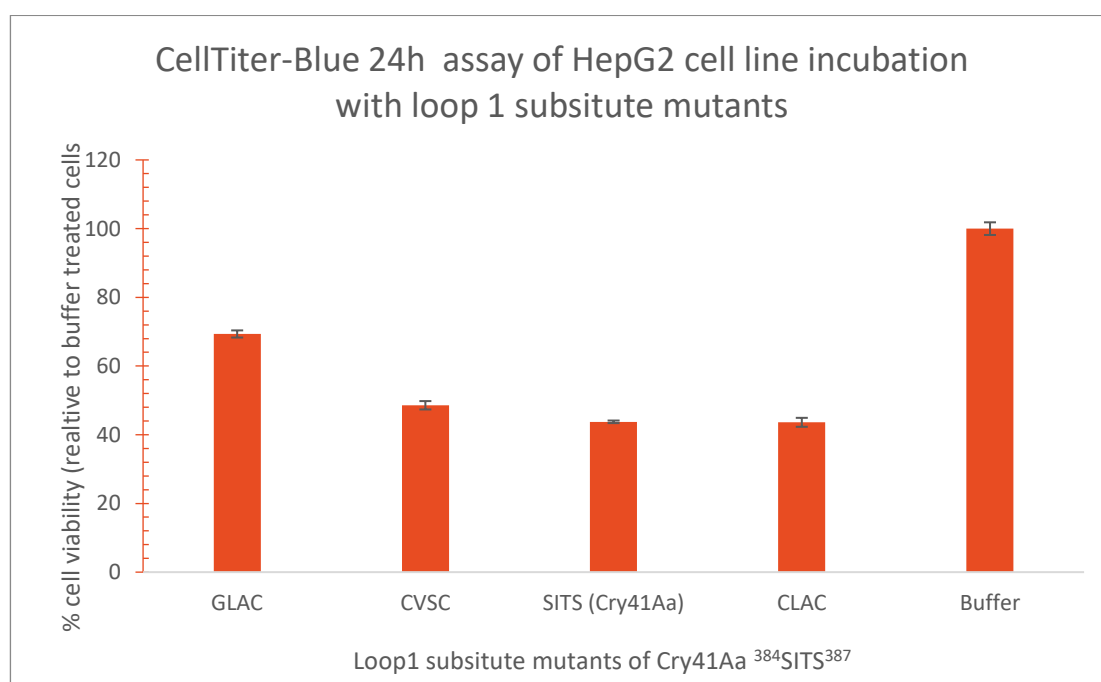
Figure 49 CellTiter blue HepG2 cell assay of loop 1 mutants
Figure shows % cell viability of HepG2 after a 24h incubation period with dialysed substitution mutants' of loop1 Cry41Aa. Protein concentrations were not standardised. HepG2 cells seeded at density 25×10^4 cells/mL.

Loop1 recombinants did not reveal many details on the specificity of Cry41Aa, the distinctive extra loop was explored next

6.4 Residue substitutions in extra loop of Cry41Aa.

Homology protein structure prediction software highlighted the distinctive extra loop which is unique to Cry41Aa. Structurally, the loop appears to extend away from the main Cry41Aa structure according to SWISS-MODEL software predication; or it is tucked towards the main structure according to Phyre software (figure 46). It has the sequence ^{261}NVSDYSRYPWTQYNQSGGFSYREAKGEYRGT^{291}.

Previous work has shown that complete or partial deletions on the loop did not result in a protease resistant core (Banani, 2013). A study by Howlader *et al.* (2009) created alanine substitutions in domain II loops 1, 2, and 3 of Cry4Aa. These mutants exhibited significant reduced toxicity to the susceptible *Culex pipinen* mosquito as well as highlighting loop 2 as a key loop in the role of receptor recognition (Howlader *et al.*, 2009). Further investigation in loop 2 of Cry4Aa led to the creation of additional mutants in which multiple residues in loop 2 were replaced with alanine resulting in significant reduction in the level of mosquitocidal activity. The findings led to the speculation that the receptor binding site (S) of Cry4Aa is different from loops 1, 2, and 3 or that there may be multiple binding sites that work cooperatively for receptor binding. (Howlader *et al.*, 2010). The role of the extra loop in Cry41Aa was investigated for specificity.

The sequence for the extra loop was analysed independently from the rest of the Cry41Aa sequence by PSIRED. The software data suggests that the extra loop has a secondary structure where residues ^{271}TQY^{273} and ^{279}FSYRE^{283} are sheets. An alanine substitution of residues ^{280}SYRE^{283} was carried out. This substitution created an alanine cassette changing the native secondary structure of the extra loop from sheets to helixes. The PCR primers used are listed in table 21. The mutagenesis design is illustrated in figure 50.

PCR primer list to create construct

Template DNA	Primer name	Oligonucleotide/ Primer sequence	5'-PHO
pBS41Aa	extra Loop JMF	5' GCTGCAGCCAAAGGAGAGTACCG 3'	Yes
	extra Loop JMR	5' GGCGGCAAAACCACCCGATTGATTATAC 3'	Yes

Table 21 PCR primer list to create construct
Table lists PCR primers used to create extra loop substituation.

The PCR product was purified and allowed to self-ligate overnight followed by a transformation into *E. coli.* The construct was processed in much the same way as that employed for all loop constructs. Schematic diagram in figure 13 in methods 3.2 details the steps and process taken to obtain extra loop mutant protein.

Native extra **loop sequence**

N V S D Y S R Y P W T Q Y N Q S G G F S Y R E A K G E Y R G T

5'AAC GTA TCC GAC TAT AGT CGT TAT CCA TGG ACA CAG TAT AAT CAA TCG GGT GGT TTT AGC TAC CGT GAA GCC AAA GGA GAG TAC CGA GGT ACA 3'
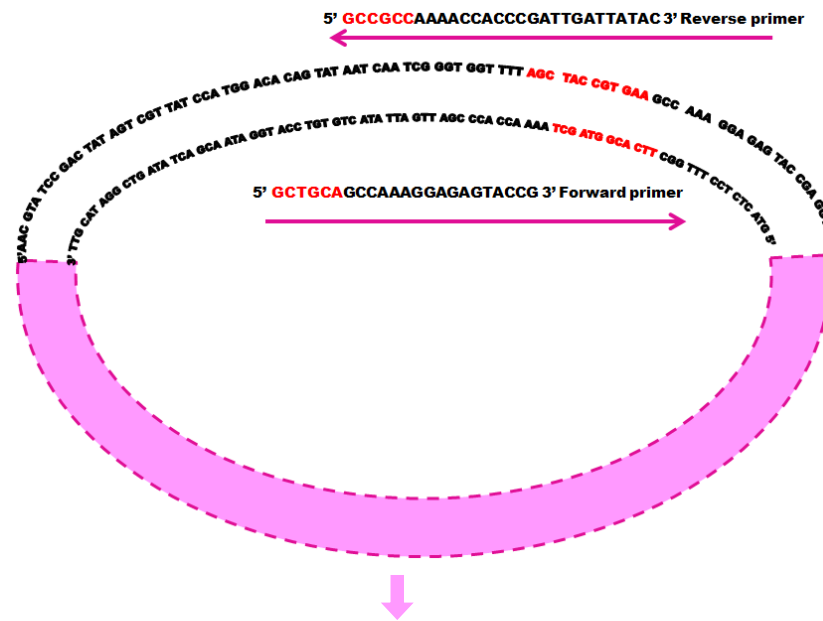
Mutant extra **loop: alanine substitutions**

N V S D Y S R Y P W T Q Y N Q S G G F A A A A A K G E Y R G T

5'AAC GTA TCC GAC TAT AGT CGT TAT CCA TGG ACA CAG TAT AAT CAA TCG GGT GGT TTT GCC GCC GCT GCA GCC AAA GGA GAG TAC CGA GGT ACA 3'

Figure 50 Schematic diagram of Primer design for Alanine substitution in extra **loop of Cry41Aa.** Substitution Residues and codon are highlighted in red in the native, primers and mutant sequence.

Colonies with potential constructs were picked and grown in ampicillin LB broth for 2-3 h. The suspensions were centrifuged, and pellet were mini prepped and to extract potential constructs. The potential constructs were profiled by Hae III digests. NEB web *cutter* software predicted the Hae III digest profiles of wildtype pBS41Aa as well as the extra loop construct.

In preparation for subcloning into a *Bt* expression vector, the constructs in blue script plasmid and the *Bt* expression vector pSVP2741Aa underwent a double digest by BamHI and Xhol. In the constructs the mutant ORF2 2.5Kb fragments was isolated. In pSVP2741Aa the backbone 7.8Kb fragment with ORF3 was isolated. Once excised, and gel purified the linear fragments were allowed to ligate overnight before an *E. coli JM109* transformation.

The transformation mixtures were plated on LB agar plates prepared with 5 µg/mL chloramphenicol and 100 µg/mL ampicillin for overnight incubation at 37 °C. The pSVP2741Aa backbone fragment encodes ampicillin and chloramphenicol genes, correctly ligated constructs ensure bacterial cell survival. The *JM109* colonies were picked and incubated for 2-3 h in antibiotic LB broth. The suspension was centrifuged, and pellets resuspended for plasmid(s) extraction.

 Hae III digests were performed on recovered mini preps. Table 22 lists the NEB web *cutter* predicted fragment sizes (bp) for the extra loop wildtype pSVP2741Aa.

Figure 51 is the gel picture of Hae III digest profile of wildtype pSVP2741Aa and the extra loop construct in *Bt* expression vector. The extra loop construct Hae III digest profile is indistinguishable from pSVP2741Aa.

Once confirmed by sequencing, the construct was introduced into *E. coli strain GM*2163 in preparation for transformation with *Bt4D7*. This removed any methylated DNA. The integrity of the construct was checked as it proceeded through various bacterial transformations via Hae III digests profiles and sequencing.

The construct was introduced into *Bt4D7* to express recombinant crystals. The transformation mixture was incubated overnight on 5 µg/mL chloramphenicol LB agar plates at 30 $^{\circ}$C. *Bt* colonies were picked and incubated in 3mLs of antibiotic LB broth for 2-3 h. The solution was centrifuged, and pellets were resuspended in order to extract the construct plasmids. Mini preps from *Bt* cells would also include native *Bt* plasmids, therefore a Hae III digest to check the correct sequence of the construct cannot be carried directly on mini preps from *Bt* cells. The *Bt* mini preps were first introduced into *E. coli* and plated onto antibiotics LB agar plates. The lawn(s) of cells were scraped, resuspended and mini prepped for the construct. A Hae III digest was performed on the recovered plasmids.

After a 3-day incubation, *Bt* cells with extra loop construct were observed for crystals under a light microscope and thereafter harvested and characterised.

Hae III digest profile by NEB web cutter

Hae III fragments	Wildtype pSVP2741Aa (bp)	extra loop in Bt expression vector Construct (bp)
1	2469	2469
2	1895	1895
3	1007	1007
4	879	879
5	629	629
6	598	598
7	587	587
8	558	458
9	458	434
10	434	306
11	306	267

Table 22 Hae III digest profile by NEB web cutter
Table lists the key fragments of the Hae III digest profile for wiltype pSVP2741Aa and the extra loop construct in *Bt* expression vector
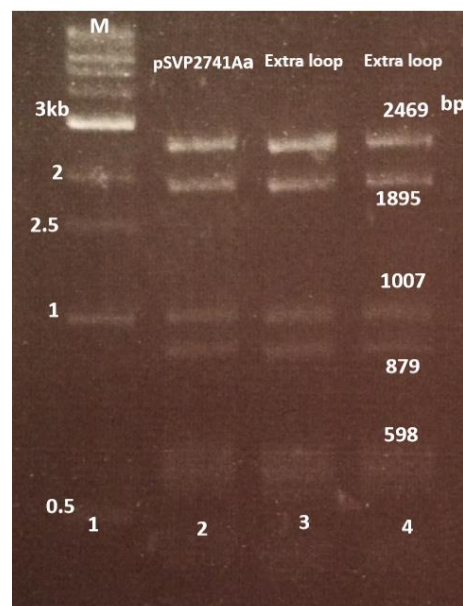
Figure 51 Agarose 1% gel of Hae III digest profile of extra **loop construct**
Lane was loaded with DNA 1kb marker. Lane 2 was loaded with wildtype Cry41Aa. Lane 3 and 4 were loaded with extra loop construct.

The extra loop mutant was solubilised in carbonate buffer pH 10.5 and activated in 1 mg/mL of carbonate trypsin. Two protease resistant proteins were observed. An ~80 KDa or slightly less as indicated with yellow arrows. A ~60 KDa proteins was also observed but it was very faint. it is indicated in with a yellow arrow in the Cry41Aa control sample. The extra loop recombinant was processed in both the 1 and 2 step method discussed in chapter 6.2.1. as indicated in the SDS PAGE gel the 2-step method was more efficient and a clear protease resistant protein in seen for recombinant in lane 3 of figure 52.
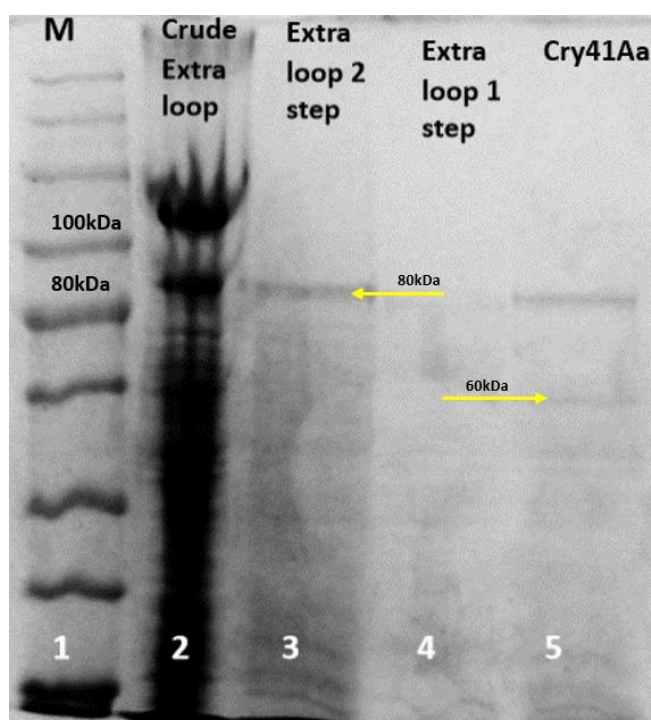
Figure 52 SDS PAGE gel analysis of the extra **loop recombinant protein against wildtype Cry41Aa protein.**
Lane 1 was loaded with protein marker. Lane 2 was loaded with 5µL of crude extra loop recombinant protein. Lane 3 was loaded with 5µL of extra loop recombinant which was processed in the 2-step method. Lane 4 was loaded with 5µL of extra loop recombinant protein processed in the 1 step method. Lane 5 was loaded with 5µL of wildtype Cry41Aa protein control.

The recombinant protein was gel purified and its concentration optimised against Cry41Aa and loop 3 recombinant proteins (chapter 6.4). A concentration range was made, and the extra loop recombinant was analysed for its effect on HepG2 cell lines. The final concentration of the extra loop recombinant in the well with HepG2 cells was 0.197, 0.131, 0.0657, 0.0263, 0,013, 0.006, 0.002 µM/ml.   A gel picture of the protein concentration optimised activated extra loop recombinant protein (0.197 µM/mL) is shown in chapter 6.5.1. figure 73. HepG2 cell assay demonstrated that this mutant lost its toxicity towards HepG2 cells as indicated in figure 74 where the graph details a dose response of HepG2 treated with all recombinant proteins whose protein concentrations were optimised.

6.5 Residue substitutions in loop 3 of Cry41Aa

A combination of bioinformatic tools predicted loop 3 of Cry41Aa (chapter 4). Mutagenesis took place on three hydrophobic residues of loop 3 at positions F509, W511, and Y514 shown in figure 53.

A number of studies have noted the likeness between the receptor binding loops of domain II and other known protein to protein epitopes. Protein to protein epitopes involve interactions between hydrophobic residues that able to bind tightly to receptors that are surrounded by hydrophobic or charged residues (Schnepf et al., 1998). In a study that examined the role of domain II, a number of alanine substitutions were made in loop 2 residues of 3-domain insecticidal CryIAb. The recombinant protein

demonstrated loss of toxicity to susceptible *Manduca sexta* and *Heliothis virescens*. The loss of toxicity directly correlated to reduced binding affinity for brush-border membrane vesicles (BBMV) prepared from susceptible insect midguts. The study proposed that ^{368}RRP^{370} residues play a key role in toxin-receptor interactions and may even direct the toxin to receptor molecule. Further investigation in the same loop analysed the role of single hydrophobic aromatic Phenylalanine residue at position 371 in the loop. A number of substitutions with hydrophilic, aliphatic residues with a small side chain were made. Observations indicated that irreversible binding and the toxicity of Cry1Ab to BBMV of *M. sexta* was significantly affect and the study concluded that a hydrophobic aromatic side-chain residue at position 371 was crucial for irreversible binding of Cry1Ab toxin to susceptible insect BBMV.

Further investigations on the effect of residue substitution in loop of Cry1Ab revealed that alanine substitutions not only affected toxicity but also binding affinity to toxin to BBMV prepared from susceptible *H. virescens* midguts (Rajamohan *et al.*, 1996b). Domain II loops of Cry1Aa toxins demonstrated that hydrophobic residues are critical for binding to BBMV of *Bombyx mori* midgut (Lu *et al.*, 1994).

Wu and Dean studied the effect of alanine substitutions in loop 1 and 3 of insecticidal Cry3Aa. The recombinants were analysed for their toxicity to the susceptible *Tenebrio molitor*. Substitutions of tyrosine residues in loop1 resulted in a reduction in receptor binding. Unexpectedly, a block substitution in loop 3 residues to alanine resulted in a reduction in receptor binding while concurrently the recombinant toxin demonstrated

an increase in toxicity by 2.4-fold when compared to wildtype Cry3Aa. Comparative analysis of the loop 3 alanine substitution block mutant and wildtype Cry3Aa conclude that mutant demonstrated increased membrane insertion into the BBMV of *T. mori* and proposed that there is a direct correlation between toxicity and irreversible binding of this mutant to susceptible BBMV. The study concluded that loop1 and 3 of Cry3Aa are involved in receptor binding and loop 3 playing a key role in membrane insertion (Wu and Dean, 1996).

Such studies have highlighted the importance of a hydrophobic residues in receptor recognition and irreversible binding to susceptible cells. In this study initial mutagenesis involved native hydrophobic, aromatic residues substitution with alanine at positions 509, 511, and 514 to create three different mutants.

6.3. Primer design for alanine substitutions at positions 509, 511, and 514 of loop 3 in Cry41Aa.

Initial mutagenesis aimed to create alanine substitution in position 509, 511, 514 of loop 3 in domain II of Cry41Aa. The three-loop 3 alanine substitution mutants F509A, W511A, and Y514A were made according to the primer design illustrated in figure 20 chapter 5.1. The wildtype pBS41Aa plasmid (figure 21) acted as a template DNA to amplify loop 3 of Cry41Aa. All three-substitution mutants F509A, W511A, and Y514A were made using pBS41Aa plasmid as a template as listed in table 23 and illustrated in figure 53.

The primers were designed so that each DNA fragment forms one half of loop 3. Once these linear PCR products were excised, gel purified and allowed to ligate, they formed a construct ready for introduction into *E. coli.* Correct ligation results in an MfeI restriction site introduced when codon change was introduced in the reverse primers. This silent mutation did not change native residue, a digest with MfeI was never carried out as Hae III digests and sequencing were sufficient in identifying constructs. The primer design is illustrated in figure 53 The constructs were introduced into *E. coli JM*109*,* and the mixture plated on 100 μg/mL ampicillin agar plates.

PCR primer list used to create constructs

Substitution	Forward Primer	PCR product sequence
F509 to A	5'GCCCTGCCGCGTGGCCTGG3'	5'GGATTAAATTTGAACCTATTAAATTTGAACCTGTACGGGAC AATTGCCCTGCCGCGTGGCCTGG3'
W511 to A	5'GCCCTTTCGCGGCGCCTGGTTATAAA3'	5'GGATTAAATTTGAACCTATTAAATTTGAACCTGTACGGGAC AATTGCCCTTTCGCGGCGCCTGGTTATAAAC3'
W511 to Y		5'GGATTAAATTTGAACCTATTAAATTTGAACCTGTACGGGAC AATTGCCCTTTCGCGTATCCTGGTTATAAAC3'
W511 to F		5'GGATTAAATTTGAACCTATTAAATTTGAACCTGTACGGGAC AATTGCCCTTTCGCGTTTCCTGGTTATAAAC3'
Y513 to A	5'GCCCTTTCGCGGCGCGCCTG GTGCTAAACAATTAAGT3'	5'GGATTAAATTTGAACCTATTAAATTTGAACCTGTACGGGAC AATTGCCCTTTCGCGTGGCCTGGTGCTAAACAATTAAGT3'
F509 to Degenerate	5' GCCCTTNBGCGTGGCCTGG 3'	5'GGATTAAATTTGAACCTATTAAATTTGAACCTGTACGGGAC AATTGCCCTTNBGCGTGGCCTGGTTATAAA3'
Reverse Primer	5'ATTGTCCCGTACAGGTTCAAATTTAATC3'	
MfeI Restriction site	5' CATTG3'	

Table 23 PCR primer list used to create constructs
Table of primers used to introduce amino acid substitution into loop 3 of Cry41Aa. Substitution amino acids are highlighted in green. A codon change allowed the introduction of an MfeI restriction site in the PCR product.
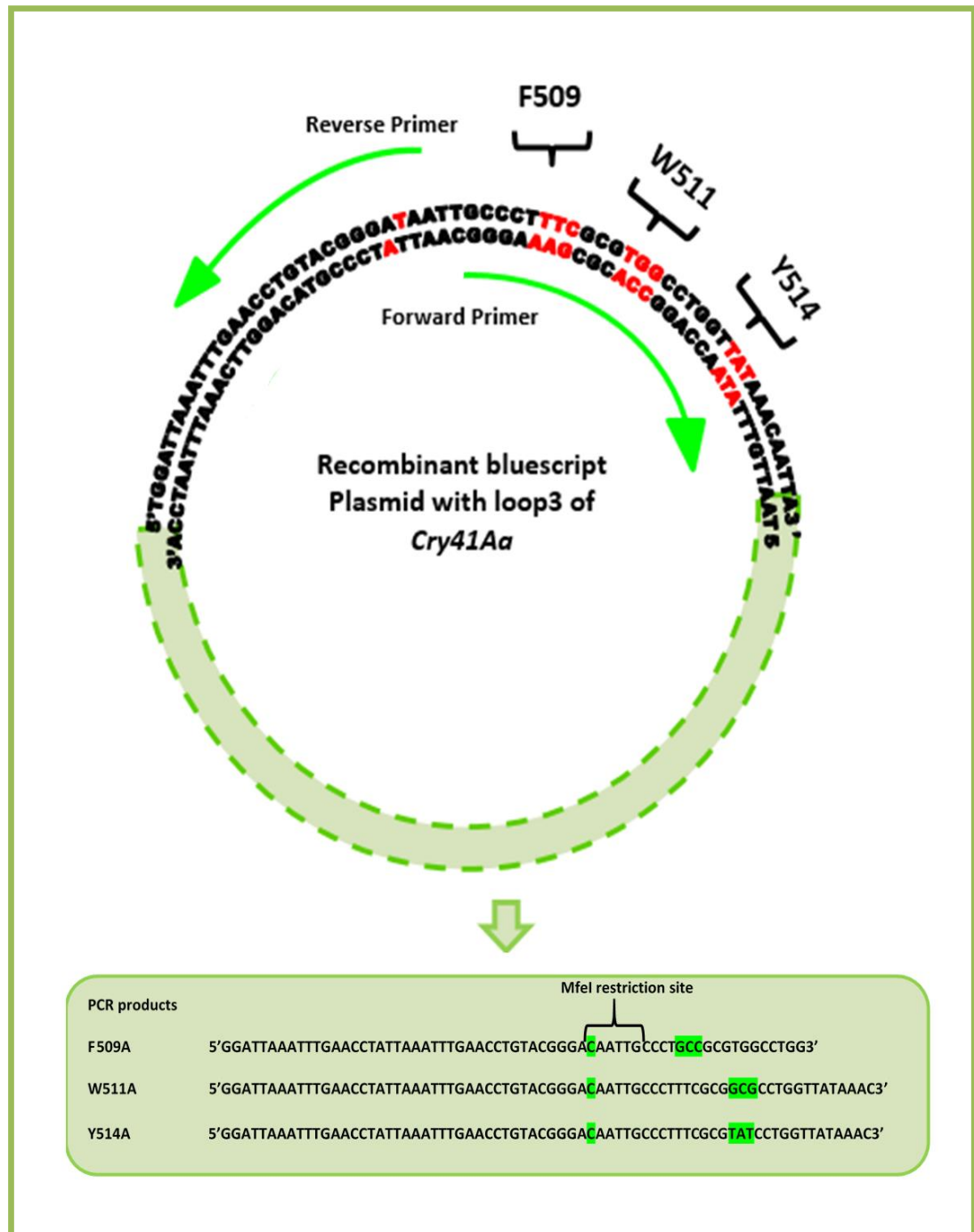
Figure 53 Schematic representation of primer design for alanine substitutions on loop 3 of Cry41Aa.
Wildtype pBS41A acted as template DNA to create construct. The position of each amino acid destined for substitution is numbered and highlighted in red in the native sequence of the plasmid. A codon change at 'T' allows the introduction of an MfeI restriction site in the PCR products.

Colonies with potential constructs were picked and grown in ampicillin LB broth for 2-3 h. The suspensions were centrifuged, and pellet were mini prepped and to extract potential constructs. The potential constructs were profiled by Hae III digests. NEB web cutter software predicted the Hae III digest profiles of wildtype pBS41Aa as well as the constructs. The Hae III profile of each mutant loop 3 varied dependents on the type of substitution made.

The construct W511A is used here as an example to illustrate the steps and check points that all constructs undertook. Table 24 lists some the predicted Hae III fragments of construct W511 and wildtype pBS41Aa as calculated by NEB web cutter.

Hae III digest profile by *NEB web* cutter

Hae III fragments	Wildtype pBS41Aa (bp)	pBS41Aaloop3 Construct (bp)
1	767	767
2	629	705
3	558	629
4	549	549
5	458	458
6	434	434
7	340	340
8	306	306
9	267	267
10	254	254

Table 24 Hae III digest profile by NEB web cutter
 lists the key fragments of the Hae III digest profile for wiltype pBS41Aa and the construct W511A in blue script

Figure 54 is the gel picture of mutant W511A which had a distinctive 705bp band the

wildtype pBS41Aa band 558bp was noticeably missing. The W511A as well as other loop

3 constructs were first identified by Hae III digests performed on mini preps of plasmids

extracted from *E. coli   JM109* cells. The constructs were subsequently confirmed by

sequencing before preparation for subcloning into a *Bt* shuttle vector.
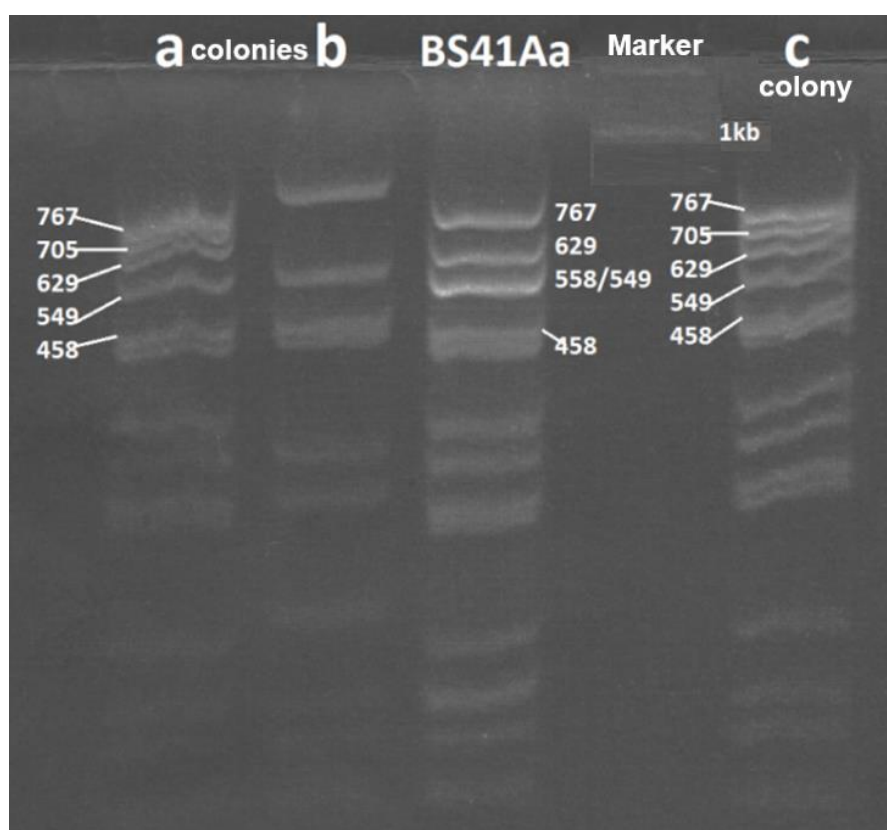
Figure 54 Banding profie of construct in agrose gel
 Agarose 1.5% gel of Hae III digest of plasmids extracted from *JM109* colonies a lane 1, colony b in lane 2 , and colony c in lane 5 .
lane 3 contained wildtype pBS41AA. Lane 4 contained DNA marker.   All lanes labelled with content.

In preparation for subcloning into a *Bt* expression vector, the constructs and the Bt

expression vector pSVP2741Aa underwent a double digest by BamHI and Xhol. In the

constructs the mutant ORF2 2.5Kb fragments was isolated. In pSVP2741Aa the backbone

7.8Kb fragment with ORF3 was isolated. Once excised, and gel purified the linear fragments were allowed to ligate overnight before a transformation into *E. coli JM*109. The transformation mixtures were plated on LB agar plates prepared with 5 µg/mL chloramphenicol and 100 µg/mL ampicillin for overnight incubation at 37°C.

The pSVP2741Aa backbone fragment encodes ampicillin and chloramphenicol genes, correctly ligated constructs ensure bacterial cell survival. The *JM109* colonies were picked and incubated for 2-3 h in antibiotic LB broth. The suspension was centrifuged, and pellets resuspended for plasmid(s) extraction. Hae III digests were performed on recovered mini preps.

Table 25 lists the NEB web cutter predicted fragment sizes (bp) for the construct W511A and wildtype pSVP2741Aa. Figure 55 is the gel picture of Hae III digest profile of wildtype pSVP2741Aa and the W511A construct in *Bt* expression vector. Colony 1 and 5 both show the distinct 705bp band of a Hae III digest profile on construct W511A. Once confirmed by sequencing, all constructs were introduced into *E. coli* strain *GM*2163 in preparation for *Bt4D7*. Transformation.

The integrity of all constructs was checked as they proceeded through various transformations. Figure 55 confirmed that the construct W511A of colony 1 and 5 were indeed the same construct that was recovered after *GM*2613 transformations. All loop 3 constructs were processed and checked in the same manner as construct W511A.

Hae III digest profile by *NEB web* cutter

Hae III fragments	Wildtype pSVP2741Aa (bp)	W511A in Bt expression vector Construct (bp)
1	2469	2469
2	1895	1895
3	1007	1007
4	879	879
5	629	705
6	598	629
7	587	598
8	558	587
9	458	458
10	434	434
11	306	306

Table 25 Hae III digest profile by NEB web cutter
lists the key fragments of the Hae III digest profile for wiltype pSVP2741Aa and the construct W511A in Bt expression vector
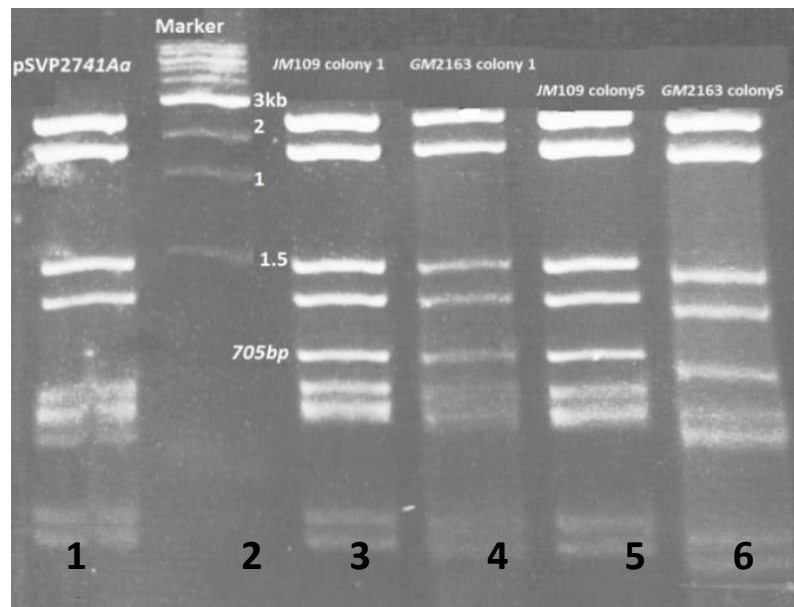
Figure 55 Banding profie of construct in agrose gel
Agarose 1.5% gel of Hae III digest performed on plasmids recoverd from *E. coli* strains *JM109* and *GM*2163 colonies. Lane 1 containe wildtype pSVP2741AA. Lane 2 DNA marker. Lane 3 conatined colony 1 extracted from JM109. Lane 4 conatined colony 1 extracted from GM2163. Lane 5 conayined colony 5extracted from JM109. Lane 6 conatined coloy 5 extracted from GM2163.
All lanes labelled with contents

Sequencing confirmed all constructs were recovered from *GM*2163 cells. The constructs were introduced into *Bt4D7* to express recombinant crystals. The transformation mixtures were incubated overnight on 5 µg/mL chloramphenicol LB agar plates at 30$^\text{o}$C. *Bt* colonies were picked and incubated in 3mLs of antibiotic LB broth for 2-3 h.

The 2mL solution were evenly distributed on 5 µg/mL chloramphenicol LB agar plates at 30$^\text{o}$C for a 3-day incubation. The remaining solution was centrifuged, and pellets were resuspended in order to extract the construct plasmids. The *Bt* mini preps were first introduced into *E. coli and* plated onto antibiotics LB agar plates. The lawn of cells was scraped, resuspended and mini prepped for the constructs. Hae III digest were performed on the recovered plasmids to check for the correct constructs.

Figure 56 is the gel picture of various Hae III digests performed on construct W511A from where it was first created in *JM109* cells to its GM2163 transformation to remove methylated DNA in order to introduce it with *Bt* cell where the recombinant crystal was expressed, to finally out of *Bt* cells and *JM109* transform to confirm the integrity of construct W511A.

All constructs were subject to Hae III digests and sequencing checks. After a 3-day incubation, samples were observed for crystals under a light microscope and thereafter harvested and characterised.
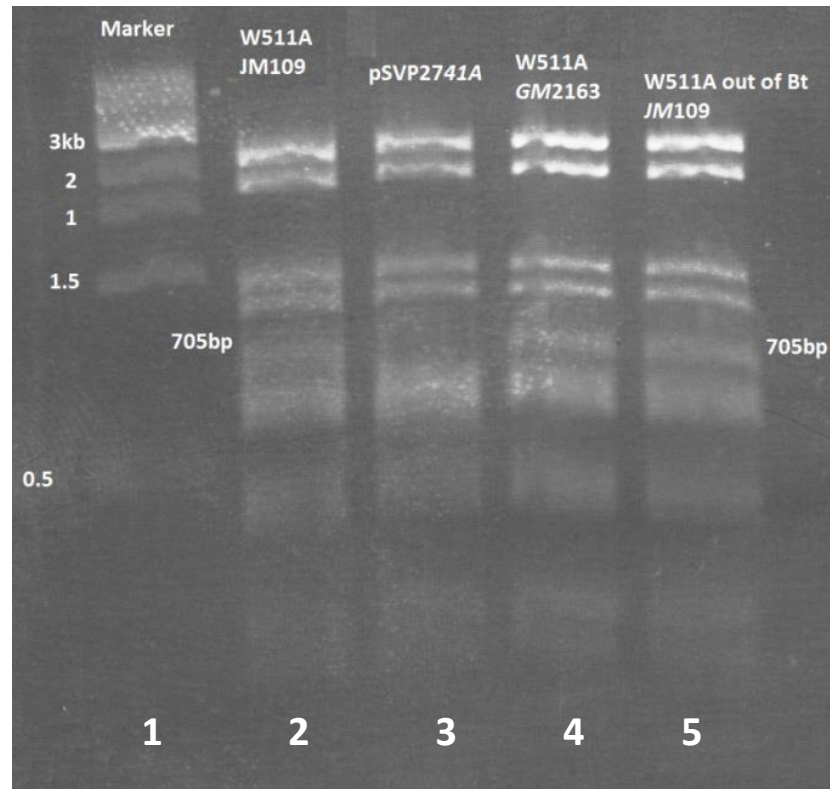
Figure 56 Banding profie of construct in agrose gel

1.5%agarose gel of various Hae III digests performed on construct W511A as it journed to and out various *E. coli* and Bt cells. Lane 1 contanined DNA marker. Lane 2 conatined W511A extracted from JM109. Lane 3 conatined wiltype pSVP2741Aa. Lane 4 conatine W511A extraced from GM2163. Lane 5 conatined W511A extracted from *JM109* post tranfromation with *Bt4D7*. All lanes labelled with content.

6.5.1 Characterisation and purification of alanine substitutes F509A, W511A, and Y514A

Bt bacterial cells undergo a stationary phase during their cycle that is thought to be triggered by a shortage in nutrients. It is during this phase that protein molecules accumulate in the mother cell as crystal inclusions constitute 25% of the dry weight of the sporulated cells. Studies have indicated that much of the *Bt* cell's energy is used for

crystal and spore production which are physiological changes indicative of the stationary phase in the bacterium's life cycle (Agaisse and Lereclus, 1995; Palma *et al* ., 2014).

Once the constructs were introduced into *Bt4D7* and plated onto 5 µg/mL chloramphenicol agar, they were incubated for 3days at 30˚C. Samples were observed under the light microscope to confirm the presence of crystals. Electron microscopy images were taken of some mutant crystal and compared to wildtype Cry41Aa crystal made *Bt4D7* strain. It shows typical bipyramidal *Bt* crystals mixed with spores (figure 57). The mutant crystals are indistinguishable from wildtype.
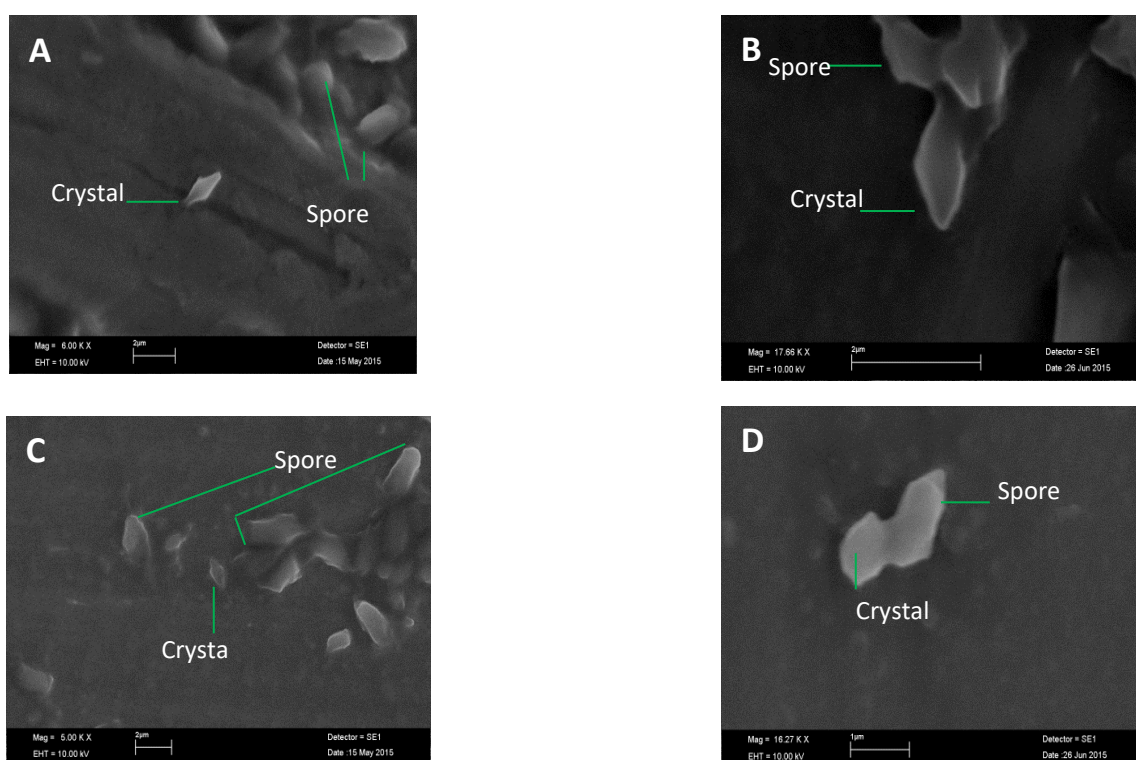
Figure 57 Electron microscope images of *Bt* crystals taken after three days of incubation during the sporulation stage of *Bacillus thuringiensis* 4D7.
Crystals and spore labelled accordingly indicated in green. Picture A=Cry41Aa (wildtype), B= F509A, C= W511A, D= Y514.

Cry41Aa wildtype and its recombinant were subjected to the two-step procedure. Figure 58 is the SDS PAGE gel picture of crude wildtype Cry41Aa and F509A and W511A loop 3 substitution recombinant crude samples. The blue box highlights the ˜120 KDa protein expressed by ORF3 gene. The yellow box highlights the ˜80-88 KDa recombinant protein expressed by ORF2 gene. All crystals made by the loop substitution mutants were observed under light microscope and then visualised in an SDS PAGE gel. All recombinant crude samples are comparatively indistinguishable from Cry41Aa crude crystals.   The loop substitute recombinants were solubilised and activated in the established 2 step process.
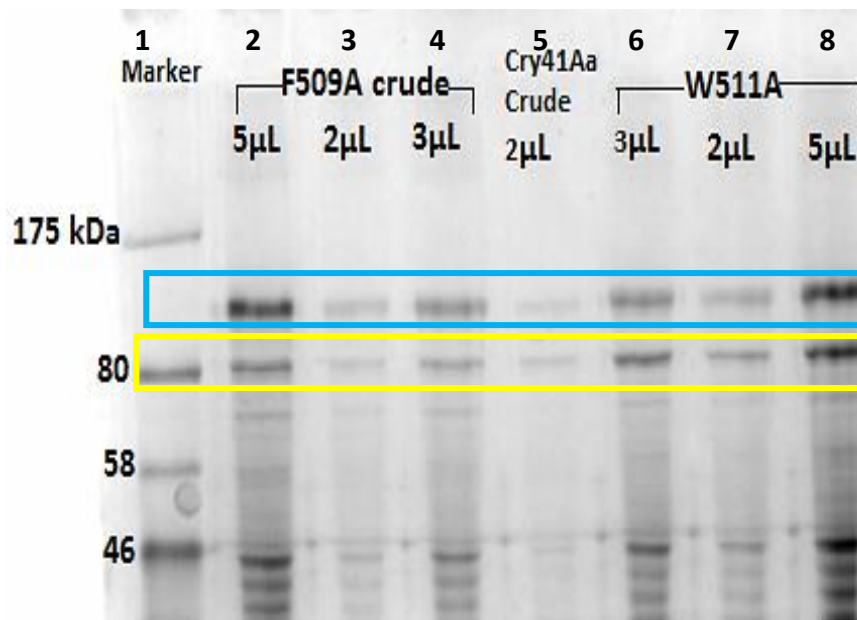
Figure 58 7.5% SDS PAGE gel showing total crude sample of recombinant F509A and W511A of different volumes against crude wildtype Cry41Aa.
The blue box highlights the 120 KDa recombinant protein expressed by ORF3 gene. The yellow box highlights the 80 KDa recombinant protein expressed by ORF2 gene. All lanes labelled with content and volume (µL)

Figure 59 shows the SDS PAGE gel analysis of the solubilised and activated F509A mutant

and Cry41Aa. The F509A mutant exhibited a faint band that represented ORF3 (~120

KDa) and an ORF2 (~80-88 KDa) in only the carbonate lane 3. Its equivalent wildtype

Cry41Aa carbonate sample was quite faint but an ORF3 protein was apparent.

Both Cry41Aa and F509A proteins showed better yield of activated toxins when

solubilisation in carbonate and DTT as indicated by prominent bands that represented

~80-88 KDa protoxin protein 5 and 6. With the exception of sample in lane 3 there was

an absence of the ~120 KDa protein encoded by ORF3 in other samples in the gel.

According to Yamashita *et al*. (2005) the ORF2 gene encodes the active toxin, thus the

88 KDa proteins is activated by proteinase K to produce a protease resistant core of ~64

KDa. Here, Cry41Aa was activated by 1 mg/mL trypsin in carbonate and produced a

prominent major protein of ~80 KDa (or slightly less) as observed in lanes 2, 7, 8, 9, and

10. The minor ~60 KDa protein appeared fainter and although it was observed in gel it

does not show well in picture.

The smaller protein~60 KDa is less concentrated than the bigger 80 KDa protein. The

activated F509A mutant resulted in a major ~80 KDa band as seen in lanes 8 and 10 of

the gel picture. The minor ~60 KDa band was also present but very faint and clear in

image. In Cry41Aa activated protein sample (lane 7 and 9) and of a similar concentration

to that of mutant F509A the minor protein is also very faint. Both F509A samples are shown in the lanes 9 and 10 of the gel picture. Protease inhibitors were added post activation to reduce degradation of protein by native proteases.
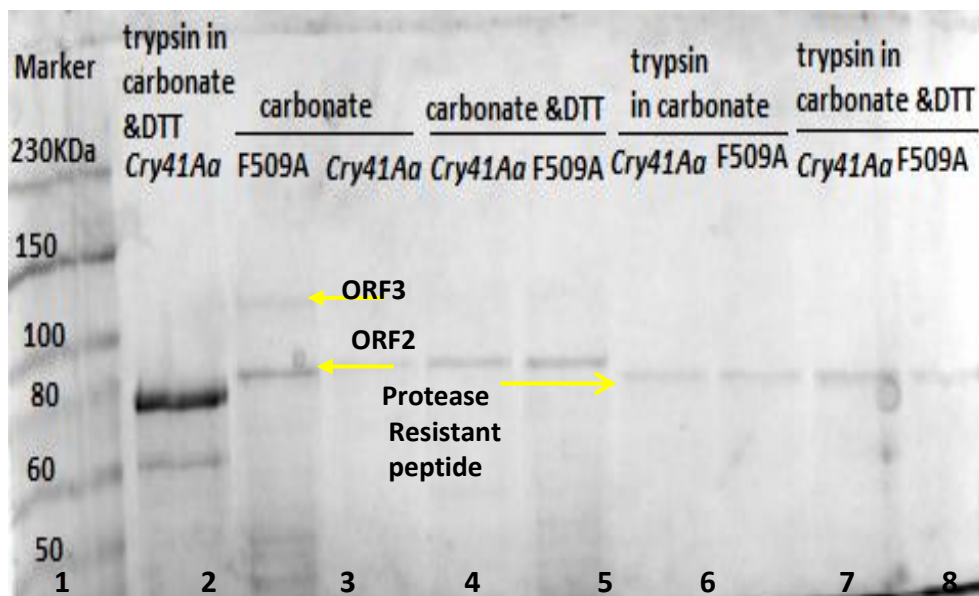
Figure 59 7.5% SDS PAGE gel of Cry41Aa and recombinant F509A in different conditions.
Lane 1 is loaded with protein marker. Lane 2 is loaded with supernatant Cry41Aa after it was solubilised in carbonate pH 10.5and DTT (2.5mM). Lane 3 was loaded with supernatant F509A after incubation in carbonate pH 10.5. Lane 4 was loaded with supernatant Cry41Aa after incubation in carbonate pH 10.5. Lane 5 was loaded with supernatant Cry41Aa after it was solubilised in carbonate pH 10.5and DTT (2.5mM). Lane 6 was loaded with supernatant F509A after it was solubilised in carbonate pH 10.5and DTT (2.5mM). Lane 7 was loaded with activated Cry41Aa which was solubilised in carbonate pH 10.5. Lane 8 was loaded with activated F509A which was solubilised in carbonate pH 10.5. Lane 9 was loaded with activated Cry41Aa after it was solubilised in carbonate pH 10.5and DTT (2.5mM). Lane 10 was loaded with F509A after it was solubilised in carbonate pH 10.5and DTT (2.5mM).

6.5.3. Purification of Cry41Aa loop substitute F509A, W511A and Y514A

The activated forms of the recombinant toxins required purification. This step aimed to isolate the two peptides thought to be responsible for toxicity. The findings from this study have identified a major protein of ~ 80 KDa (or slightly less) and a minor protein of ~60 KDa.

Anion exchange chromatography was employed to isolate each protein separately, so that each could be analysed for its toxicity towards HepG2 cell lines. The pI of Cry41Aa was calculated as 6.18 by compute pI/Mw Expasy tool. Yamashita, *et al*. (2005) successfully separated individual peptides via anion exchange. Cry41Aa and its recombinants F509A, W511A, and Y514A were dialysed overnight against 10mM CAPS buffer at pH 10.4 in order to remove DTT and other low molecular weight contaminants. They were then eluted by increasing sodium chloride gradient from 0 to 1 M. The fractions were collected and visualised on SDS PAGE gels.

Activated Cry41Aa was purified using AKTA to separate the two peptides (~80 KDa and ~60 KDa) and visualised in SDS PAGE gels as shown in figure 60a. The toxin was eluted as a single peak (figure 60b), and the collected fractions always show traces of the minor 60 KDa protein particularly once the fractions were concentrated in preparation for cell assays of toxins. In this study, it was not possible to individually isolate each of the protein in the wildtype toxin. However, Domanska, 2016 did manage to separate the ~80

KDa protein and the ˜60 KDa peptides from each other by activating with proteinase K and concluded that both proteins were subject to proteolytic activation and exhibit cytotoxicity to HepG2 cells (Domanska, 2016, Souissi, 2018).

The recombinant toxins F509A, W511A, and Y514A were also AKTA purified and fractions collected as shown in figure 60-63. Similarly, to Cry41Aa gel images, the minor 60 KDa protein band is not obvious in some SDS PAGE gels of the collected AKTA purified loop recombinants fractions, but they were quite evident once the sample were concentrated in preparation for cell assay investigations.

(a)

Marker dialysed filtered A9 A10 A11 A12 B12 B11 B10 B9 B8 B7 B6 — Cry41Aa

80

60kDa

(b)

Peak

Concentration of eluted peptide

Low   High
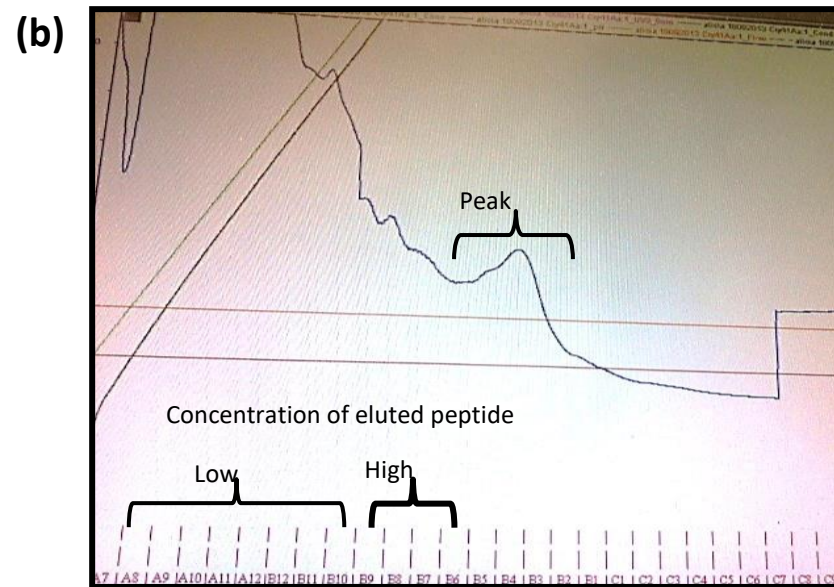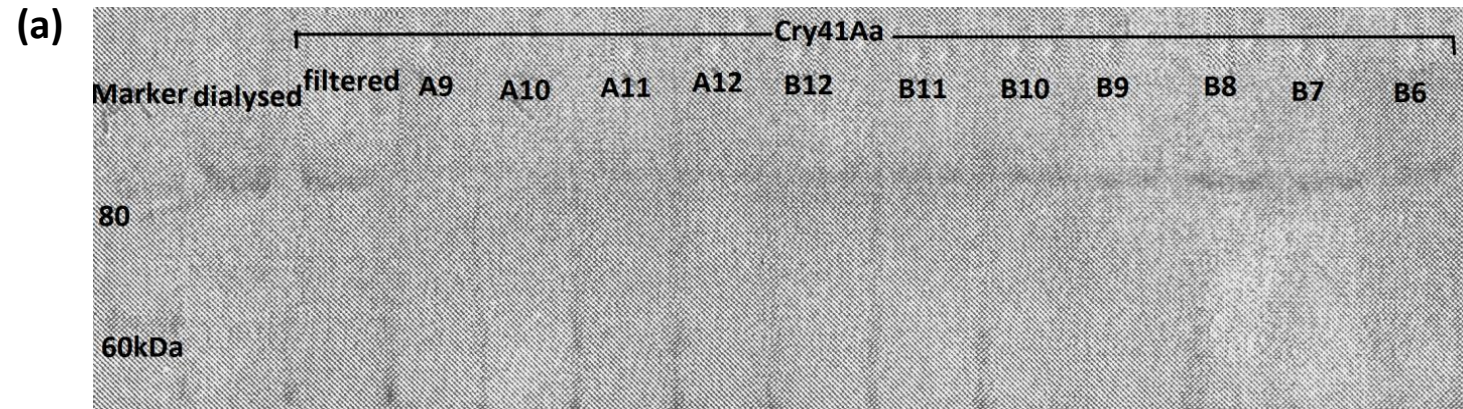
Figure 60 of AKTA purified fractions of activated Cry41Aa.
(a) SDS PAGE 7.5% gel of AKTA purified fractions of activated Cry41Aa. (b) AKTA elution profile of trypsin activated Cry41Aa. Toxin was eluted in 10mM CAPS, pH 10.4 buffer as a single peak at
~750mM sodium chloride

(a)
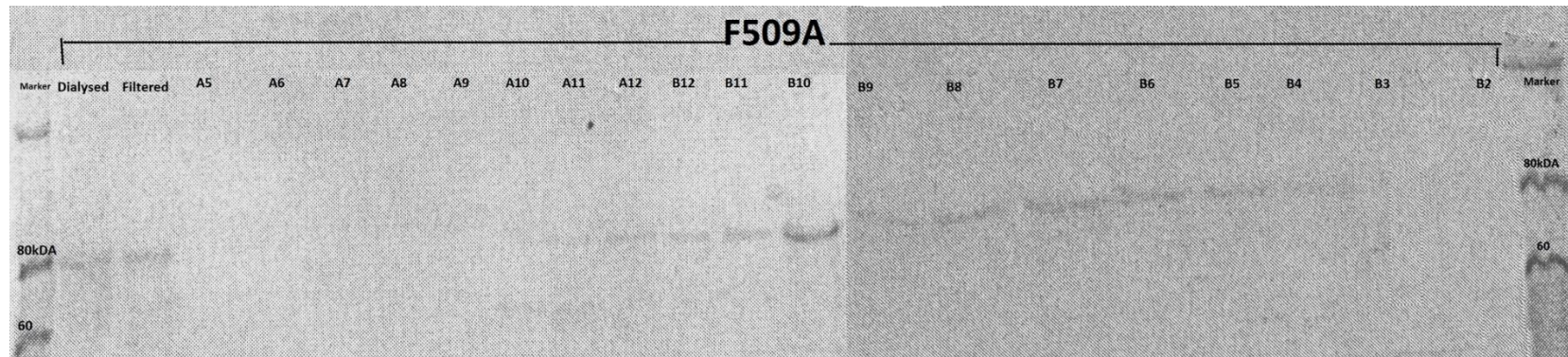
(b)
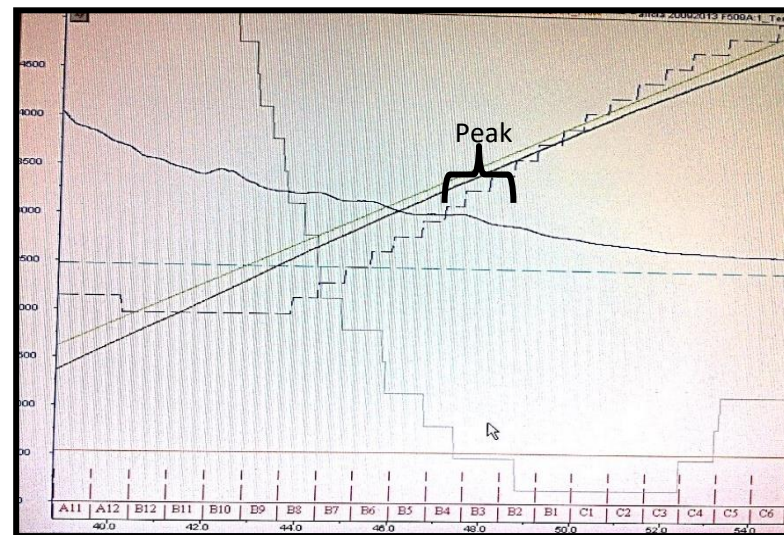
Figure 61 AKTA purified fractions of activated recombinant F509A
(a) SDS PAGE 7.5% gel of AKTA purified fractions of activated recombinant F509A.
(b) AKTA elution profile of trypsin activated recombinant F509A. Toxin was eluted in 10mM CAPS, pH 10.4 buffer as a single peak at ~750mM sodium chloride.
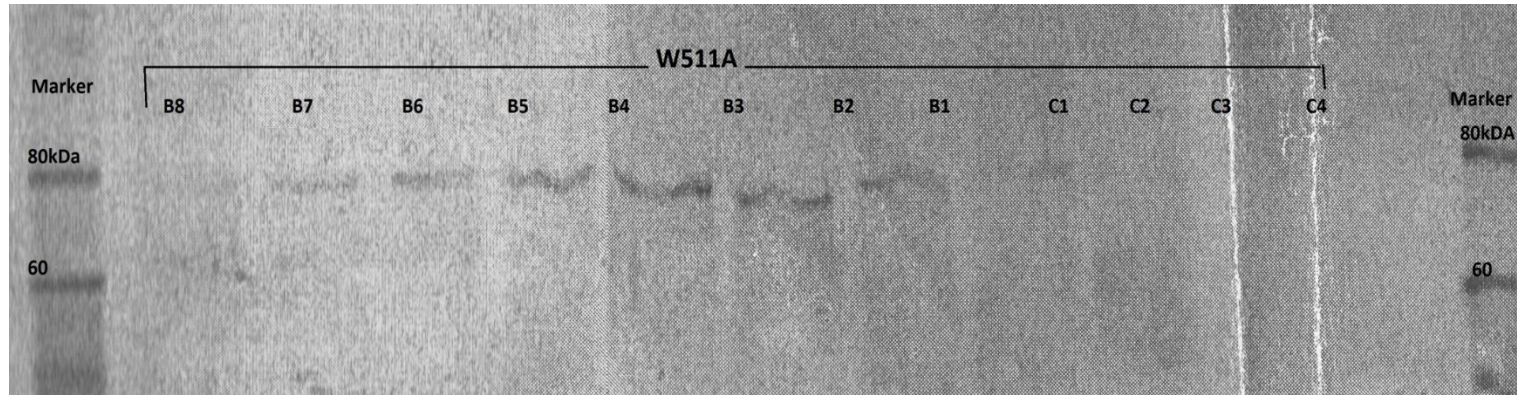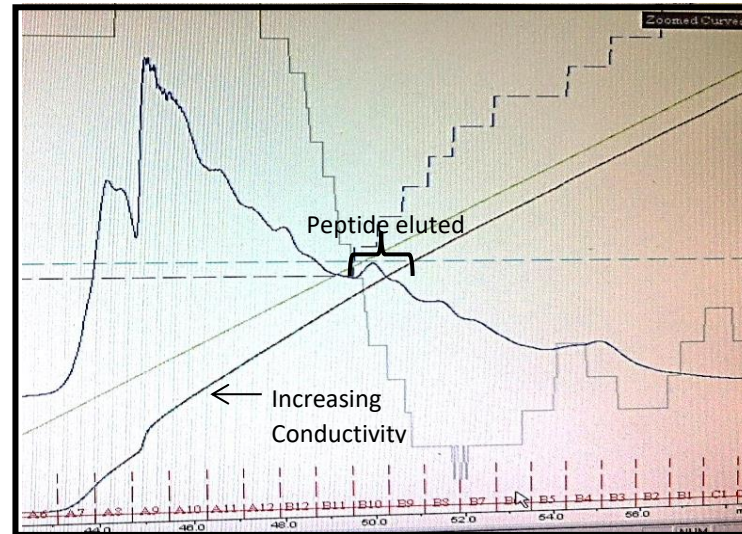
Figure 62 AKTA purified fractions of activated recombinant W511A
(a) SDS PAGE 7.5% gel of AKTA purified fractions of activated recombinant W511A. (b) AKTA elution profile of trypsin activated recombinant W511A. Toxin was eluted in 10mM CAPS, pH 10.4 buffer as a single peak at ~750mM sodium chloride.

(a)
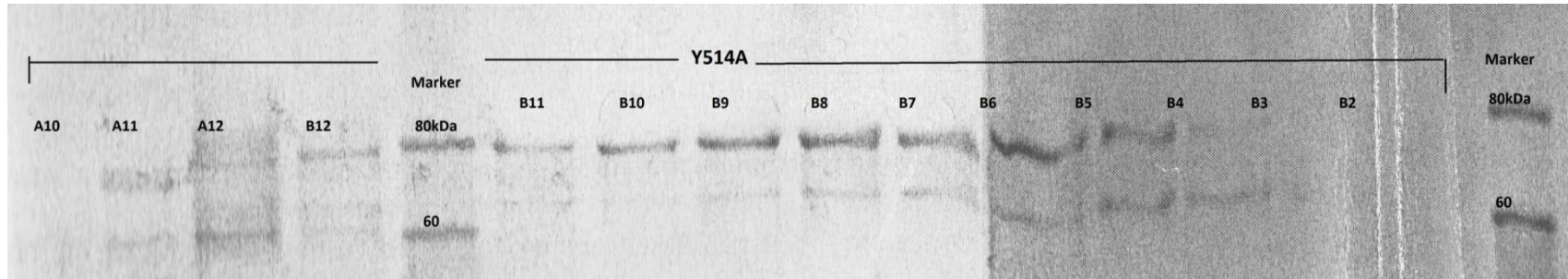
(b)
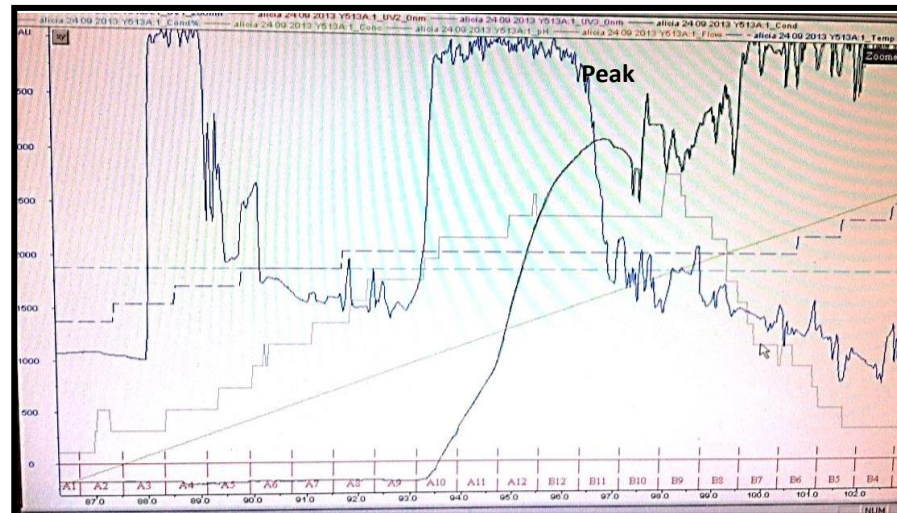
Figure 63 AKTA purified fractions of activated recombinant Y514A

(a) SDS PAGE 7.5% gel of AKTA purified fractions of activated recombinant Y514A. (b) AKTA elution profile of trypsin activated recombinant Y415A. Toxin was eluted in 10mM CAPS, pH 10.4 buffer as a single peak at ~750mM sodium chloride

.

6.5.4. Establishing cytotoxicity of F509A, W511A, and Y514A recombinants proteins

Preliminary experiments were carried to establish the cytotoxicity status of Cry41Aa toxin and its mutants towards HepG2 cell line. Cell assays of HepG2 cell incubated with AKTA purified recombinant proteins were carried out. Figure 64 is the gel image of AKTA purified Cry41Aa, F509A, W511A, and Y514A. showing activted protein of ~80 KDa of similar concentration Table 26 lists the spectroscopic analysis from Braford protein assay carried out on preliminary AKTA purified Cry41Aa, F509A, W511a, and Y514A.
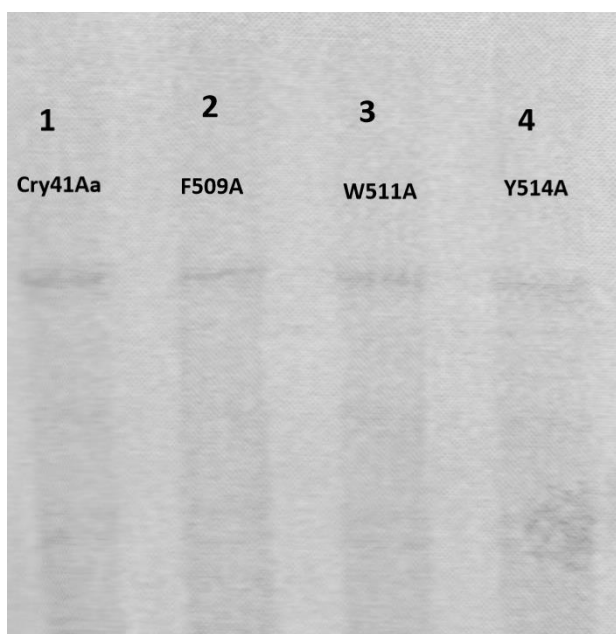
Figure 64 SDS PAGE gel of purified recombinant samples used on HepG2 cell assay
The protein concentration is optimised to approximately 100µg/mL. lane 1 was loaded with 5µL of Cry41Aa. lane 2 loaded was loaded with 5µL of recombinant protein F509A. Lane 3 was loaded with 5µL of recombinant W511A. lane 4 was loaded with 5µL of recombinant Y514A. the samples were used in HepG2 cell assay.

Protein concentration assay

Toxin name	protein concentration (mg/mL)
Cry41Aa	0.487
F509A	0.463
W511A	0.481
Y514A	0.484

Table 26 Protein concentration assay
Table lists the spectroscopic analysis from Braford protein assay carried out on preliminary AKTA purified Cry41Aa, F509A, W511A, and Y514A

The toxins were incubated with HepG2 cell lines for 24 h, thereafter the cell viability was measured using CellTiter-Blue assay. The assay estimates metabolic activity of cells using a fluorometric method. The resazurin dye is nontoxic but permeable to cells, once it enters viable cells it is reduced to high fluorescent resorufin, the signal is measured to give an estimate of viable cells (O'Brien *et al*., 2000).

CellTiter-Blue assay was added as an end point reagent of toxin-cell experiments and incubated for two hours. Hela and HepG2 cells were seeded at different densities in a 96 well plate at 22500 cells/well and 5000. Hela is a cancerous cell line that is not susceptible to Cry41Aa and thus acted as negative control to Cry41Aa susceptible HepG2 cell line (Yamashita., 2005). The Crickmore lab had previously carried out cell assays investigations on Cry41Aa and had used the 5000 cell/well density as a standard (Krishnan, 2013). The higher cell density of 2×10^4 was used by Yamashita *et al.* (2005) and was used here to check the consistency of the toxicity of Cry41Aa.

Cry1Ca is well characterised insecticidal 3-domain toxin which was expressed in *Bt4D7* strain in the same manner as Cry41Aa. HepG2 cells are not susceptible to Cry1Ca toxin and it therefore acts a negative control for 3-domain toxins made in the lab in *Bt4D7*. The HepG2 cell lines were incubated with buffer used to dialyse the recombinants as well as TX-100 (detergent that disrupts cell membrane) and etoposide (cytotoxin). TX-100 caused cell death by necrosis and etoposide by apoptosis. These were used as positive controls for toxicity i.e. cell death.

Figure 65 is the graphical presentation of a preliminary investigation on different cell densities of both HepG2 and Hela cells incubation with Cry41Aa (wildtype), the recombinant toxins F509A, W511A, and Y514A, as well as the positive cell death controls TX-100 and etoposide. At this stage the toxins were of similar concentration.

The data indicated that Cry41Aa and its three loop 3 recombinant mutants F509A, W511A, and Y514A do not have any effect on cancerous HeLa cell at 22,500 cells per well. Unfortunately, a cell assay with HeLa cells at a lower density was not carried out. It cannot be stated that the loop 3 recombinant proteins or Cry41Aa are not toxic to Hela cells at a lower cell density. Cell assay studies have demonstrated that toxins can exert a different toxicity depending on cell densities used (Soberón *et al.,* 2018).

The cell assay registered almost 0 % cell viability with TX-100 and a 75 % cell viability after etoposide incubation. This observation has highlighted the specificity of the Cry41Aa toxin and to an extent the specificity of its recombinants to HepG2 cells.

The Y514A mutant has retained toxicity to HepG2 cells with a 20% cell viability in higher cell density compared to 40% in lower cell density. Wildtype Cry41Aa toxins was also toxic to HepG2 cell lines but not as fatal as recombinant toxin Y514A with 38 % cell viability (higher cell density) and a 60 % cell viability in lower cell density. Microscopic observations of HepG2 cell lines in incubation separately with Cry41Aa and Y514A toxins indicated that Y514A induced immediate cellular swelling and appeared to be more potent than wildtype Cry41Aa. It was too soon at this stage to recognise Y514A as more potent compare to the wildtype Cry41Aa. The protein concentrations required optimisation and visualisation on an SDS PAGE gel first. A dose response experiment would also confirm if Y514A was indeed more toxic than wildtype Cry41Aa.

Both F509A and W511A toxins were not toxic to HepG2 cell line, along with insecticidal Cry1Ca both these recombinants had no effect on the viability of HepG2 cells. Under the microscope, HepG2 cells appeared intact and viable. This loss of toxicity was evident. Thus, substitution in both recombinant toxins may have affected how they interact with HepG2 cells which resulted in the disrupted specificity or reducing cell-toxin(s) interactions. Alternatively, it is also possible that interactions did still occur but no longer led to cell death.
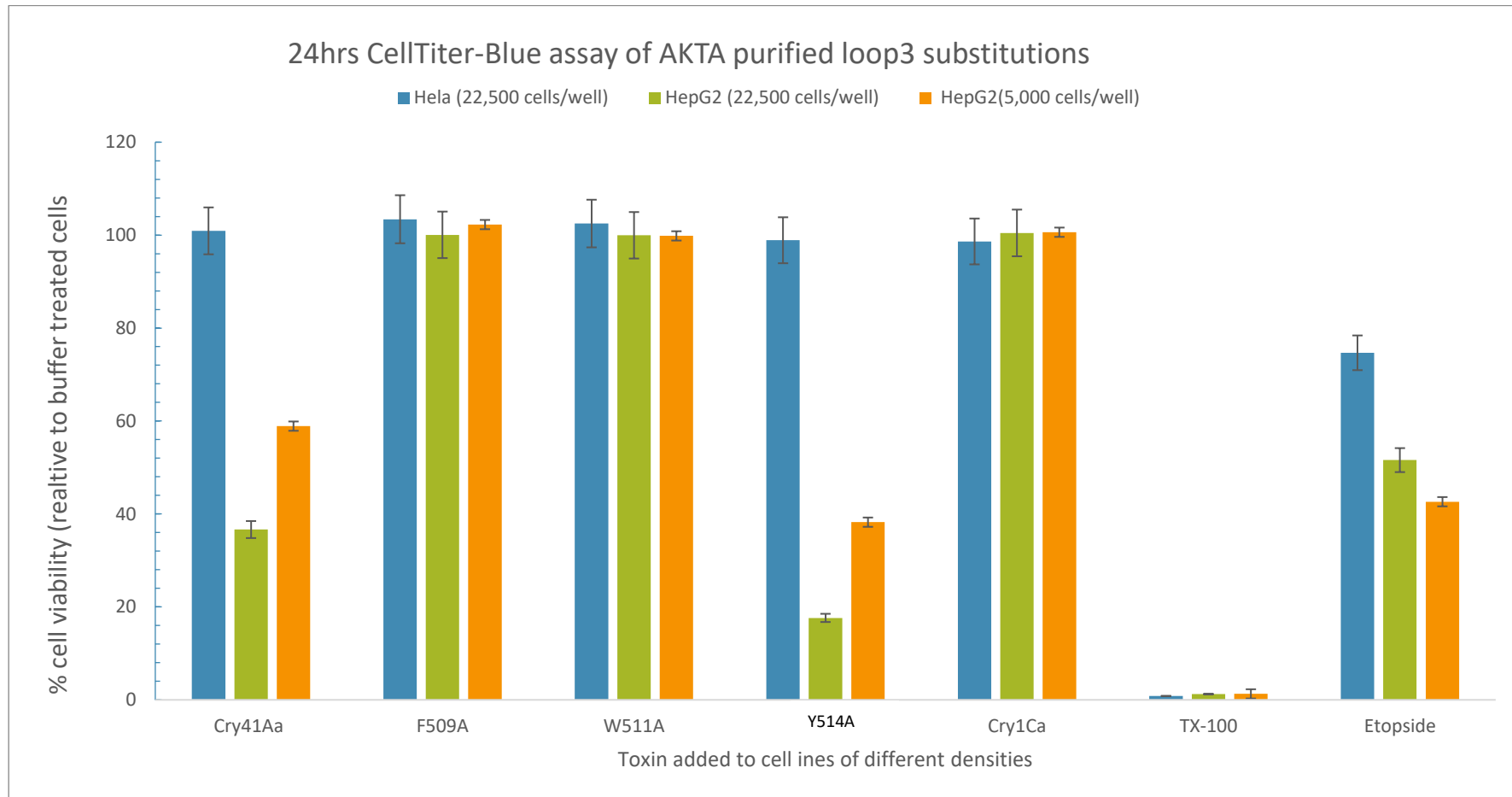
Figure 65 CellTiter-Blue 24h assay of loop 3 substitutions, F509A, W511A, and Y514A at 1.3 µM/ml on Hela (seeded at 22x10^4) and HepG2 (seeded at 5,000 and 22x10^4) cell lines at different cell densities. Cry1Ca act as negative controls. TX-100 and Etopside act as positive controls.

The effects of Cry41Aa and its loop 3 recombinant F509A were investigated for their effects on cell membrane damage with the use of cytotoxic markers such as CellTox-Green cell assay kit. Cell membrane permeability can be determined when a small fluorescent DNA binding molecule permeates the damaged membrane and binds to the cell's DNA. The fluorescent signal correlates to membrane damage and increases in a time-dependent manner.

HepG2 cells were incubated for a period of 28 hrs and cytotoxicity was measured every half an hour for the first 3hrs and then every hour for the following 3hrs with reading at 26hrs of incubation and a final reading at 28hrs of incubation. Figure 66 is the graphical representation of HepG2 cell membrane damage as a result of incubation with recombinant toxin F509A and wildtype Cry41Aa.

CellTox green fluorescent signal increased proportionally as Cry41Aa incubation time with HepG2 cell line also increased. The fluorescent signal from HepG2 cells incubated with F509A remained very low even after 28hrs of incubation and is not different to readings from negative controls buffer and insecticidal Cry1Ca. Low CellTox green fluorescent signals suggest that F509A did not induce HepG2 cell membrane damage and can be interpreted as lack of toxicity by F509A toxin. The echo findings from CellTiter-Blue analysis on F509A confirm that the single amino acid substitute mutant as nontoxic to HepG2 cell lines.
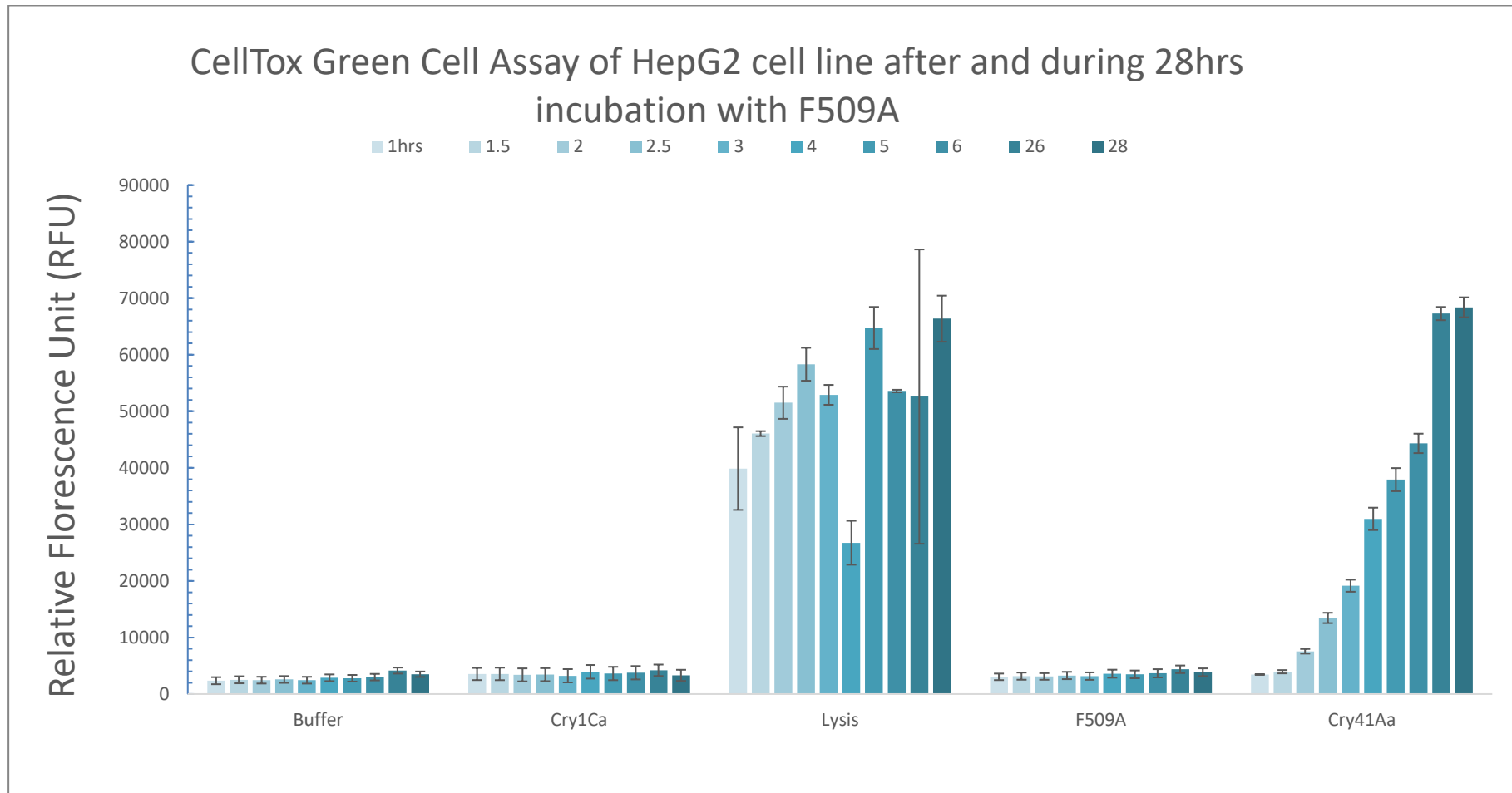
Figure 66 CellTox green assay
Membrane damage assessments by CellTox green in HepG2 cells incubated over a 28hrs period. HepG2 cells were seeded at 25x10^4 cell/mL with CellTox-Green dye in a black 96-well plate. The next day the cells were incubated with Cry41Aa, F509A, Cry1Ca at 1.3 µM/ml concentration, lysis, or buffer. The fluorescent signal was measure at various time points post toxin addition.

HL-60 was reported as a susceptible cell line to Cry41Aa (Yamashita *et al.*, 2005). In this study cell viability analysis has concluded that Cry41Aa as well as its loop mutant F509A, W511A, and Y514A, did not exhibit toxicity towards HL-60 cell lines (data not shown). The cytotoxicity of these mutants and Cry41Aa towards other human cell lines was investigated and the findings are presented in figure 67.

Therapeutic drugs can readily induce apoptosis in both Burkitt's lymphoma and non-cancerous lymphoblastoid cells lines. Thus, investigating how Cry41Aa and its loop 3 recombinants affect these cells lines can shed light on its specificity, particularly when compared to the susceptible HepG2 cell line. Cell viability was calculated in percentages relative to buffer values and determined by metabolic activity of viable cells. Table 27 lists the cell lines tested.

Figure 67 is a graphical representation of CellTiter Blue assay analysis of the cell lines listed in table 27 after a 24 h incubation with Cry41Aa F509A, W511A, and Y514A toxins. For statistical comparison of treatments in experiments, a Post-Hoc analysis applying a Bonferroni adjustment was carried out in SPSS. The cell line MUTU1 was not susceptible to any of the toxins tested. In the Y514A treatment of MUTU1 a high percentage of cell viability was recorded. It is likely that is due to human error during in the experimental procedure. The Post-Hoc analysis did not find any significance to this value when compared with buffer only treatments.

BL31 cell line treatments with toxins and buffers indicated a decrease in the percentage of cell viability. It is possible that the cell line was being affected by other factors such as contaminants or unfavourable buffer conditions. When a Post-Hoc statistical analysis was carried out there was no significant difference found between toxin treated BL31 and buffer only treated cells. Microscope observations confirmed presence of detached and ruptured cells in buffer and toxin treated BL31 wells. The cell lines GM12878 and IB4 had a 100% cell viability relative to buffer after a 24 h period of incubation with the said toxins. HepG2 cell lines were treated with toxins, only two toxins had a significant effect on percentage cell viability when compared to buffer only treated cells. Cry41Aa recorded 45 % cell viability and Y514A a 38 % cell viability after a 24h incubation period.

List of mammalian cell lines

Cell line	Origin
HepG2	Hepatocyte cancer
Hela	Uterus cervix cancer
Mutu1	Burkitt's lymphoma (EBV positive)
BL31	Burkitt's lymphoma (EBV negative)
IB4	Lymphoblastoid (B cells transfected with EBV
GM12878	Lymphoblastoid (B cells transfected with EBV)
HL-60	Acute myeloid leukaemia

Table 27 List of mammalian cell lines
Table lists the cell lines used to investigate the cytotoxicity of Cry41Aa, F509A, W511A, and Y514A.
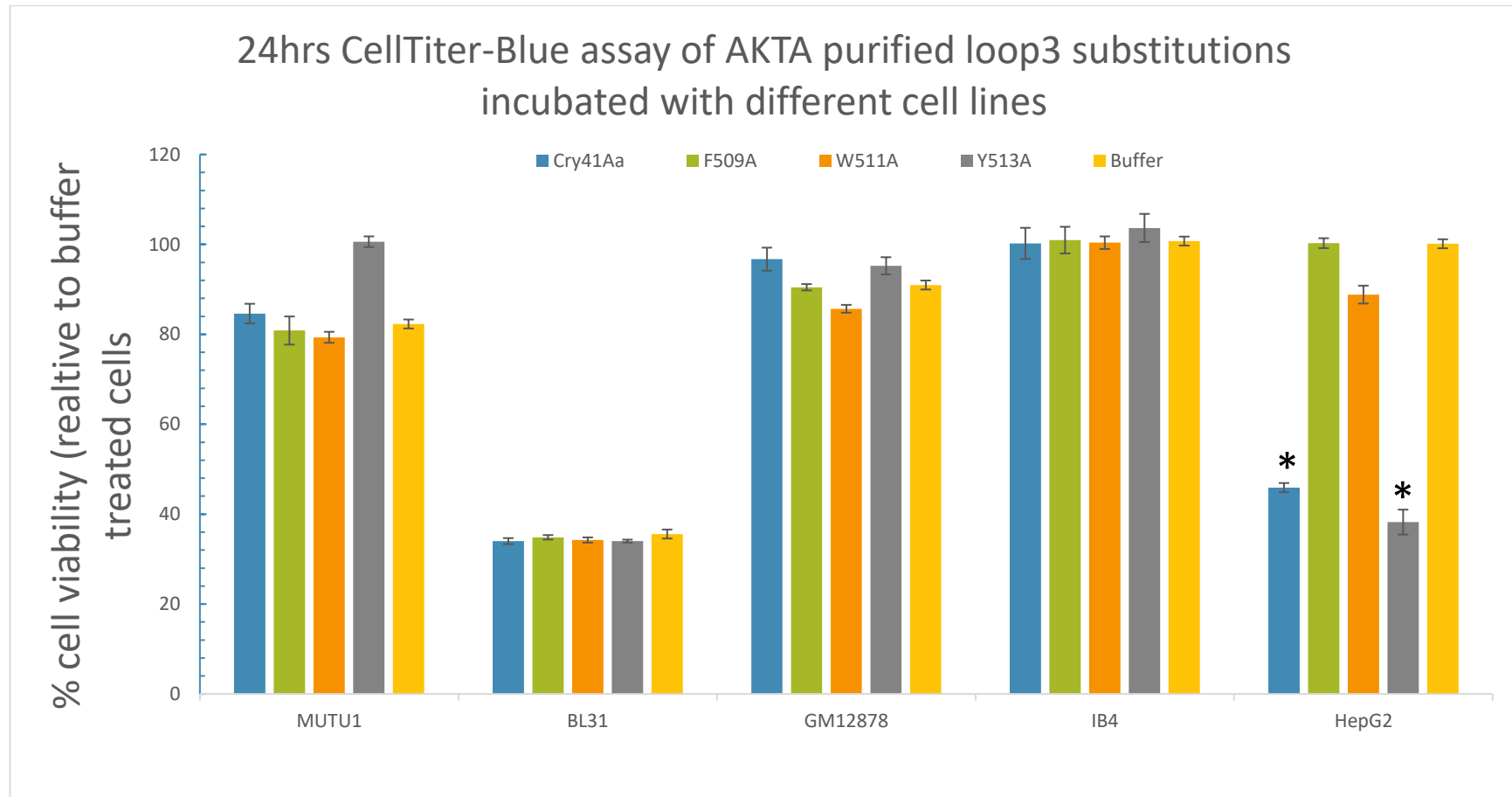
Figure 67 CellTiter-Blue Assay of various cultured cell lines.
The cells were seeded 24 h before the day of experiment at $25x10^4$ cells/mL. Cells were treated with activated Cry41Aa and loop 3 recombinant at 1.3 µM/ml. Viability was measured as a metabolic activity of viable cells 24h h post-treatment. Bars with asterisks are significantly different from buffer only control cells (* $p < 0.005$, Post-Hoc comparison with Bonferroni correction).

The lack of toxicity by F509A and W511A highlighted the need to further investigate the type of residue and the position of these residues. Based on these findings the study carried out mutagenesis on residue substitutions in positions F509 and W511. A number of degenerate substitutions were made in position F509, as well as two aromatic residue substitutions at position W511.

6.6 Degenerate amino acid substitutions at positions 509 and 511

The degenerate substitution at position 509 of Cry41Aa loop 3 was created using BS41Aa plasmid and the primers listed in table 28 below. The PCR design for the degenerate substitution is detailed in figure 68 and could generate a number of substitution combinations.

Mutagenesis carried out in loop 3

Substitution	Forward Primer	PCR product sequence
F509 to Degenerate	5' GCCCT TNB GCGTGGCCTGG3'	5'GGATTAAATTTGAACCTATTAAATTTGAACCTGTACGGGAC AATTGCCCT TNB GCGTGGCCTGGTTATAAA3'
W511 to F	5'GCCCTTTCGCG TTT CCTGGTTATA AAC3'	5'GGATTAAATTTGAACCTATTAAATTTGAACCTGTACGGGAC AATTGCCCTTTCGCG TTT CCTGGTTATAAAC3'
W511 to Y	5'GCCCTTTCGCTATCCTGGATTATA AAC3'	5'GGATTAAATTTGAACCTATTAAATTTGAACCTGTACGGGAC AATTGCCCTTTCGCG TAT CCTGGTTATAAAC3'
Reverse Primer	5'ATT G TCCCGTACAGGTTCAAATTTAATC3'	
Mfel Restriction site	5' CATTG3'	

Table 28 Mutagenesis carried out in loop 3
Table of amino acid substitutions carried out in loop 3 of Cry41Aa. Amino acids changes are highlighted in green.
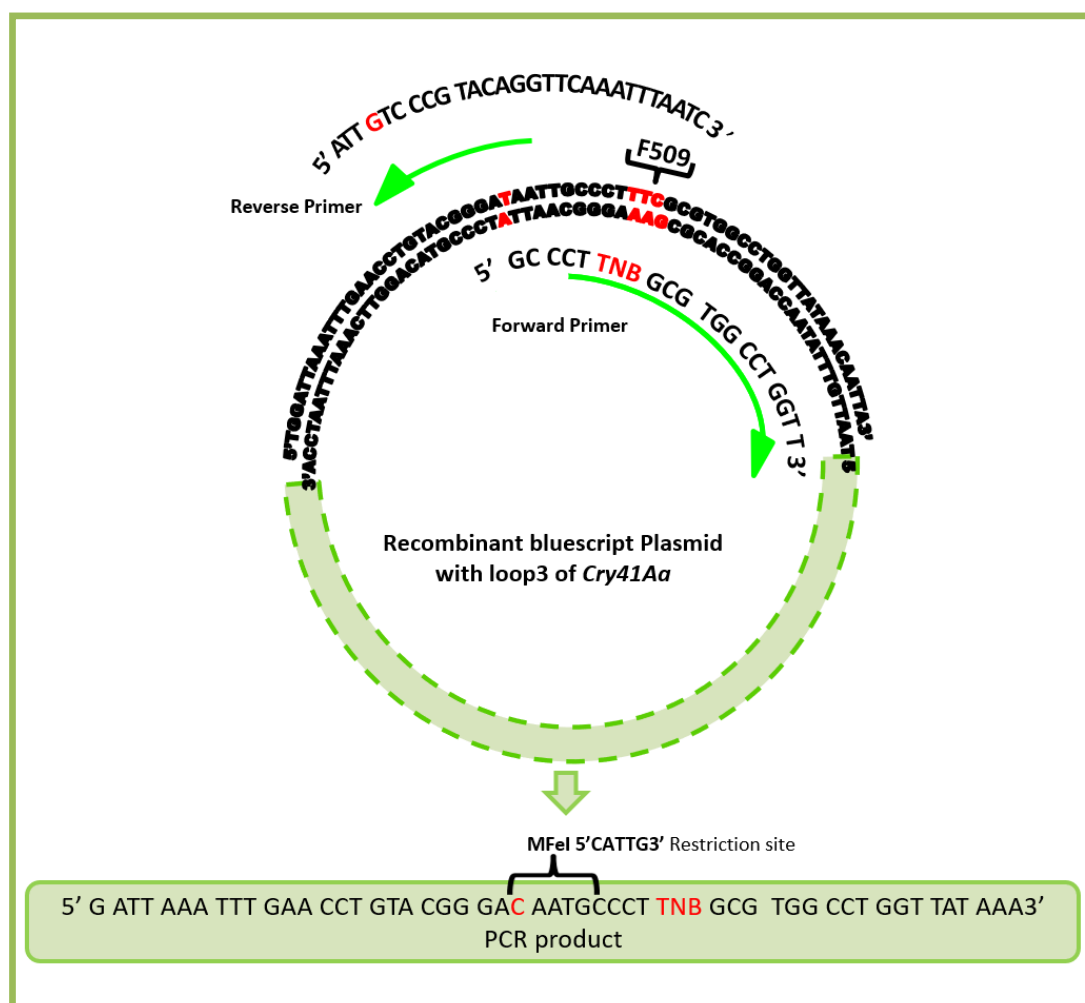
Figure 68 Schematic diagram of primer design for degenerate substitutions on loop 3 of Cry41Aa at position 509. Wildtype pBS41A acted as template DNA to create construct. The position amino acid F509 destined for substitution is numbered and highlighted in red in the native sequence of the plasmid. A codon change at 'T' allows the introduction of an MfeI restriction site in the PCR products.

The PCR product was gel purified and allowed to self-ligate overnight, before the first of many *E. coli* transformations took place. The transformation mixture was placed on two 100 μg/mL ampicillin agar plates.

Two strategies were applied to maximise the number of degenerate constructs. In the first strategy, individual colonies were picked from one ampicillin agar plate. A number

of Rapid Size Screens (RSS) of colonies as well as wildtype pBS41Aa were carried out, this gave a quick assessment of which colonies may have potential degenerate construct when compared on an agarose gel to wildtype pBS41Aa. Once the numbers of potential colonies were narrowed down, the colonies underwent a Hae III digest.

The Hae III digest profiles of some constructs were similar to that of wildtype pBS41Aa and were indistinguishable on gel from wildtype. Thus, individual degenerate substitutions were confirmed by sequencing. Once identified a double digest with BamHI and XhoI was carried out to isolate the mutant OR2(s). These were ligated overnight with pSVP2741Aa back bone fragment 7.8kb that encode ORF3 to create constructs that can be expressed in *Bt* cells.

The second strategy involved a mass scrape of colonies from the remaining plate. 1.5mL of distilled water was added to scrape any colonies, the solution was centrifuged, and the pellet was resuspended in distilled water followed by a double BamHI-XhoI digest to isolate the 2.5Kb fragments that encode the mutant ORF2(s). These were ligated over night with the 7.8kb backbone fragments of pSVP2741Aa plasmid to create a *Bt* expression vector before an *E. coli* transformation. At this stage the constructs undergo a process of selection by RSS. A construct with an RSS size that is identical on an agarose to wildtype pSVP2741Aa was selected for further analysis.

Figure 69 is the picture of agarose gel of RSS carried out in colonies with potential degenerate construct. Constructs are in *Bt* expression vector and were extracted from colonies 8 to 21. Of those colonies 10, 11, 12, 13, 14, 15, 16, 17, 20 and 21 showed an RSS size similar to wildtype pSVP2741Aa.

These were subject to Hae III digests and later confirmed by sequencing. Confirmed degenerate constructs in *Bt* expression vectors were introduced into *E. coli* GM2163 to remove methylated DNA, before a transformation with *Bt4D7* where they incubate for 3 days at 30˚C to allow for sporulation and production of crystals. Constructs were extracted from *Bt* cells for a final transformation into *E. coli* cells. This final step involved further Hae III digests to confirm the integrity of constructs.

Figure 70 is the gel picture of construct F509S and shows the Hae III digest profile of F509S where it was first identified and extracted from colony 4. The construct shares the same digest profile as wildtype pSVP2741Aa plasmid was hence it was in disguisable on gel. The construct was confirmed via sequencing. The gel pictures confirmed the construct as it journeys through different bacterial host cells.
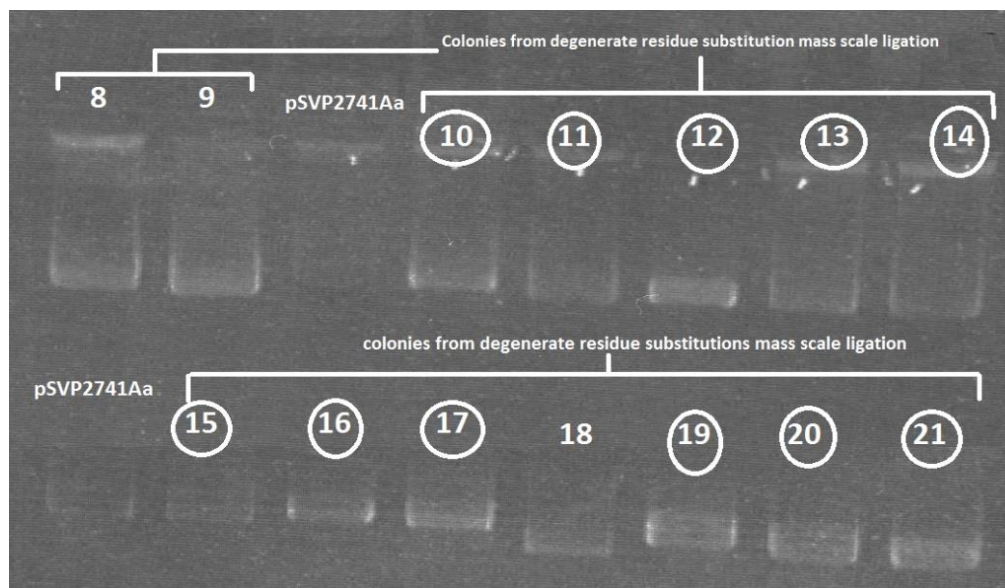
Figure 69 RSS analysis of colonies with potential construct
Double deck 1% agarose gel showing RSS on colonies 8 to 21 harbouring degenerate substitution construct after mass scale ligation of 2.5kb fragments with 7.8kb backbone of pSVP2741Aa. Upper deck was loaded with colonies 8 to 14 with Wildtype pSVP2741Aa labelled accordingly in the gel. Lower deck was loaded with colonies 15 to 21 with pSVP2741Aa labelled in lane. Colonies with potential degenerate construct and similar sizes to wildtype are circled.
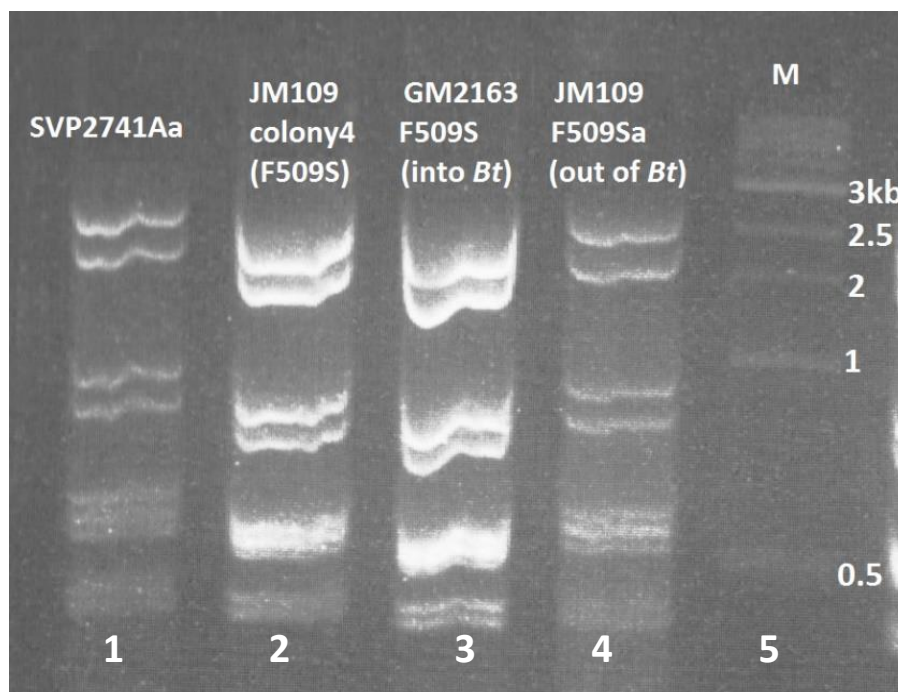

Figure 70 Banding profie of construct in agrose gel
Agarose 1.5% gel of Hae III digest profiles of construct F509S as the constructs are hosted by different *E. coli* strains lane 1 contained control pSVP2741Aa. Lane 2 contained colony 4 extracted from JM109. Lane 3 contained construct F509S extracted from GM2163 prior to transformation with *Bt4D7*. Lane 4 contained construct F509Sa extracted from *JM109* after extraction from *Bt4D7*. All lanes labelled with content.

The following loop 3 mutants were created F509A, F509Y, F509S, F509C, F509, W511A, W511Y, W511F, Y513A and their crystals were observed under light microscope. Figure 71 is the electron microscope image of some recombinant crystals.
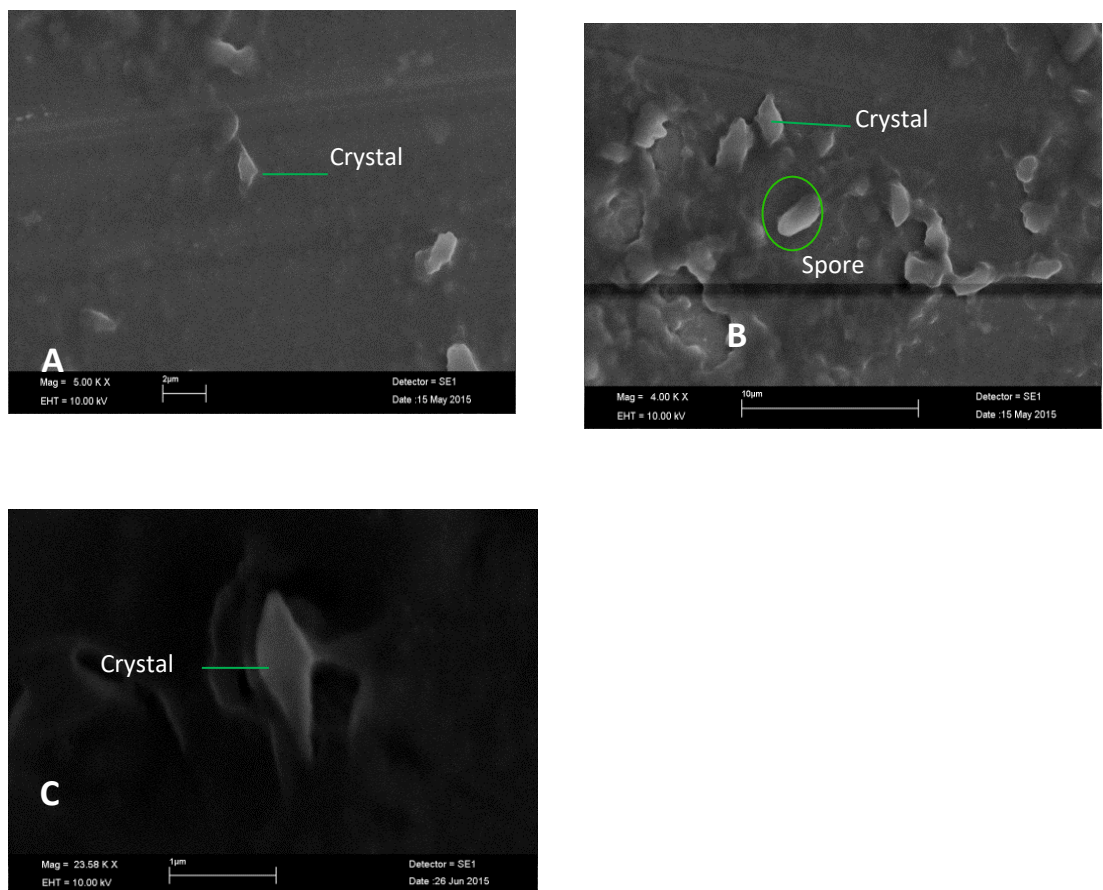

Figure 71 Electron microscope images of *Bt* crystals taken after three days of incubation during the sporulation stage of *Bacillus thuringiensis* 4D7.
Crystals and spore labelled accordingly; some spores are circled in green. Picture A=F509L, B= F509W, C= F509S.

A mass production of all loop recombinant proteins was carried out. A total of 10 mLs of crude protein of each recombinant was made and solubilised in 50mM sodium carbonate pH 10.5 and DTT (2.5mM) in two 5mLs batches. After an hour in a 37˚C water bath and occasional vortex, the samples were centrifuged, and the supernatants were recovered. Figure 72 illustrates the SDS PAGE gel analysis of the efficacy of the mass production of loop recombinants protein. Once solubilised the samples were centrifuged to separate the protoxins in the supernatant from the pellet.   The absence of proteins in lanes 2 and 3 was suggestive of the efficacy of the mass solubilisation event. The solubilised recombinant F509L and F509Y proteins in lane 4 and 5 are indistinguishable from solubilised Cry41Aa in lane 1 and 4. All recombinant proteins were subject the same solubilisation, activation and purification procedure as wildtype Cry41Aa. Due to the large volume of proteins for purification the samples were purified by column gel filtration.
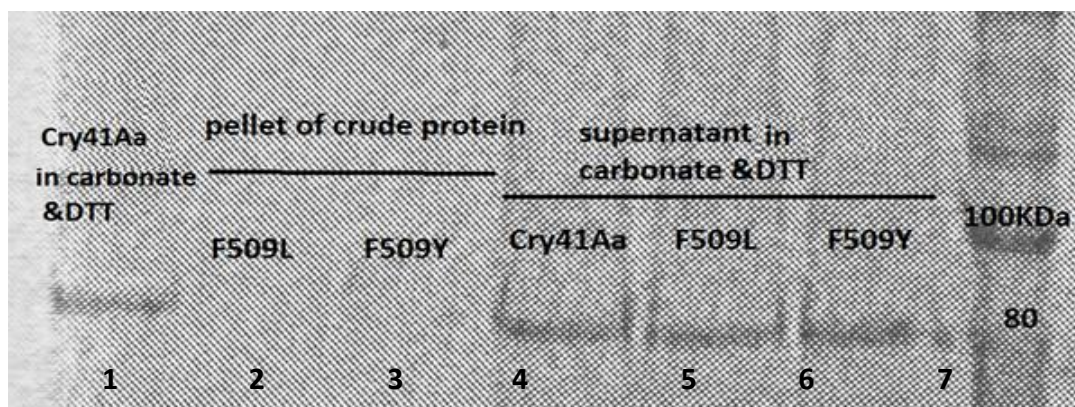
Figure 72 7.5% SDS PAGE of mass scale solubilisation of Cry41Aa, F509L and F509Y.
10mLs of crude protein of each recombinant was made and solubilised in 50mM sodium carbonate pH 10.5 and DTT (2.5mM) in two 5mLs batches in incubated in a 37°C water bath for 1 h. Centrifugation separated pellets from supernatants, and both were analysed in SDS PAGE gel. Lane 1 was loaded with solubilised Cry41Aa in carbonate pH 10.5and DTT (2.5mM). lane 2 was loaded with F509L pellet sample after supernatant was removed. Lane 3 was loaded with F509Y F509L pellet sample after supernatant was removed. Lane 4 was loaded with solubilised Cry41Aa in carbonate pH 10.5and DTT (2.5mM). lane 5 was loaded with was loaded with solubilised F509L in carbonate pH 10.5and DTT (2.5mM). lane 6 was loaded with solubilised F509Y in carbonate pH 10.5and DTT (2.5mM). lane 7 was loaded with protein marker.

A stock of 10 mg/mL of trypsin in 50mM sodium carbonate pH 10.5 was made and added to the supernatant samples to activate the recombinant toxins. This resulted in a final trypsin concentration of 1 mg/mL. Activated recombinant toxins were removed from the 37°C water bath and prepared for dialysis, purification and optimisation of the protein(s) concentration relative to each other. After activation by trypsin, protease inhibitors were added at end as it was confirmed that protease inhibitors significantly slow the rate of Cry41Aa degradation at temperature above room temperature (Domanska, 2016). Table 29 summarise the mutagenesis carried out in loop 3 of Cry41Aa.

Mutagenesis carried out in loop 3

Type of substitutions		Native residues of Cry41Aa loop 3
		^{503}V R D N C P **F** A **W** P G **Y** K Q L ^{517}
		5'GTA CGG GAT AAT TGC CCT **TTC** GCG **TGG** CCT GGT **TAT** AAA CAA TTA 3'
1st mutagenesis	Alanine	GC**C**　　　GC**G**　　　　GC**T**
2nd mutagenesis	Degenerate at position 509	**T**NB **T**AG **T**TT **T**GC **T**CC **T**CT **T**AT **T**AC **T**TG **T**GG **T**GT
3rd mutagenesis	Phenylalanine/ Tryrosine at position 511	**T**TT **T**AT

Table 29 Mutagenesis carried out in loop 3
Table of amino acid substitutions carried out in loop 3 of Cry41Aa. Native codons of amino acids target for mutagenesis are highlighted in green.

6.6.1. Optimisation of recombinant proteins concentration

In preparation for cell assay analysis, protease inhibitors were added to activate the toxins. The activated recombinant toxins were concentrated with viva spin tubes.

A stock of each was made with a concentration of 150 µg/mL according to Braford protein assay. This is equivalent to 0.197 µM/ml. The stock samples were visualised on

three SDS PAGE gels and analysed by Image Lab software. Figure 73 is the picture of one of the three SDS PAGE gels. All recombinants displayed both the major ~80 KDa protein and the minor ~60 KDa protein. The gel showed that the ~80 KDa (slightly less) bands were of identical intensity across all recombinant mutants and wildtype Cry41Aa toxin. The 60 KDa bands have more variation in their intensity and hence it was not possible to separate the two proteins and optimise the minor band intensity across all recombinants. The recombinants were similar to Cry41Aa and were indistinguishable on an SDS PAGE gel (figure 64), which suggest that they likely have a similar protein folding pattern (Krishnan *et al.*, 2017).

Image lab software was employed to analyse the band intensities and was able to calculate and quantify a band relative to a reference or a control band. The reference band was given a value of 1, other bands compared to it were given a value that was either higher or lower than 1. In this case the two reference bands are the major (~80 KDa) and minor (~60 KDa) of Cry41Aa with a concentration of 150 µg/mL or 0.197 µM/ml as determined by Braford protein assay.

Table 30 lists the values for band(s) densities for all activated recombinant toxins with a concentration of 150 µg/mL (or 0.197 µM/ml) relative to Cry41Aa bands of the same concentration. The average values of the major bands were consistent and close to the Cry41Aa control band values. The minor band(s) values showed slight variation

compared to its Cry41Aa minor control band. Hence, in this study it was not possible to separate the major and minor bands to calibrate the bands independently. The software also calculated the molecular weight of proteins based on information gathered against the protein marker of the same gel.

Table 31 shows the average molecular weights of each activated recombinant toxin and Cry41Aa toxin. The results were consistent with the control activated Cry41Aa. In addition to the following 150 µg/mL ((or 0.197 µM/ml) concentration optimised recombinant toxins F509L, F509Y, F509A, Y514A, F509W, F509S, W511A, and extra loop. The recombinant toxin W511F was also made. The concentration of this recombinant was also optimised to a concentration of 150 µg/mL according to Bradford protein assay.
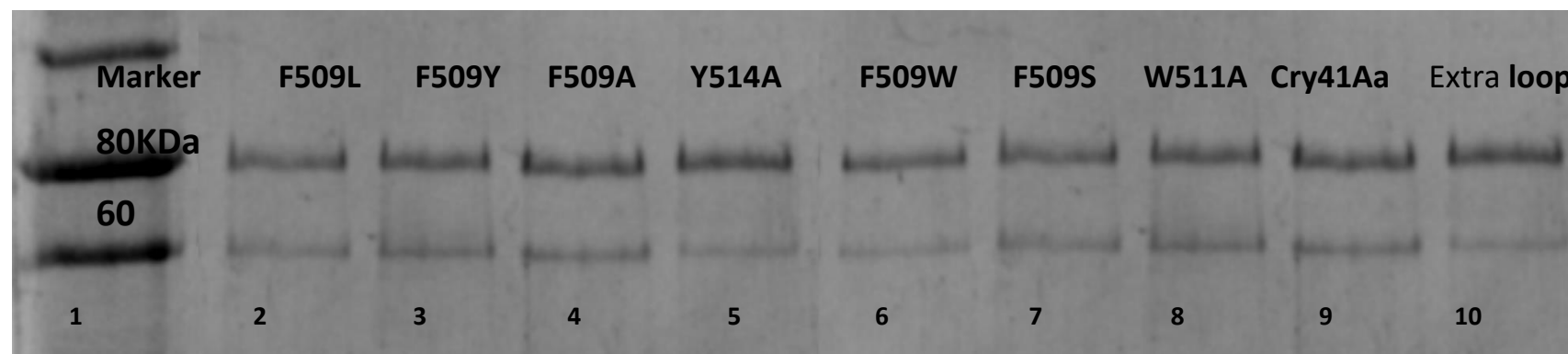
Figure 73 calibrated concentrations of activated recombinant toxins.
SDS PAGE 7.5% gel of stock samples of 150 µg/mL (or 0.197 µM/ml) of activated recombinant toxins are labelled on the gel. Cry41Aa acts as the control sample. All lanes labelled with content.

Protein concentration analysis of SDS PAGE gel with recombinant toxins

Relative qualification of recombinant toxin concentration from SDS PAGE gel bands density against Cry41Aa (150µg/mL or 0.197 µM/ml*)*

Recombinant toxin	Upper major band density				Lower minor band density			
Cry41Aa control reference	Gel 1	Gel2	Gel3	Average	Gel 1	Gel2	Gel3	Average
(150µg/mL or 0.197 µM/ml*)*	1	1	1	1	1	1	1	1
F509A	0.627491	0.799602	1.005402	0.810831667	1.169772	0.786371	0.864311	0.940151333
F509W	0.840239	1.335881	0.666376	0.947498667	0.660156	0.881126	0.724691	0.755324333
F509L	0.946215	0.811198	0.811224	0.856212333	1.123384	0.934124	0.925462	0.994323333
F509S	0.893227	0.665652	1.021243	0.860040667	1.341398	0.764116	0.844891	0.983468333
F509Y	0.736255	1.420251	0.733981	0.963495667	1.272967	0.736255	1.123241	1.044154333
W511A	1.272112	0.767586	0.772999	0.937565667	0.983826	1.180234	0.867586	1.010548667
Y514A	0.636653	0.837145	1.121344	0.865047333	0.710095	0.725442	0.813021	0.649519333
extra loop	0.992829	1.118322	0.404306	0.838485667	1.348998	0.803918	0.901101	1.018005667

Table 30 Protein concentration analysis of SDS PAGE gel with recombinant toxins
Image Lab quantification of band density of 150 µg/mL of activated recombinant toxins relative to Cry41Aa.

Molecular weight analysis of SDS PAGE gel with recombinant toxins.

Relative qualification of recombinant toxin molecular weight (KDa) from SDS PAGE gel bands

Recombinant toxin	Upper major band KDa				Lower minor band KDa			
	Gel 1	Gel2	Gel3	Average	Gel 1	Gel2	Gel3	Average
Cry41Aa	89.44272	82.50056	83.89433	85.2792	68.36135	63.29883	66.34767	66.0026
F509A	86.3987	83.78	83.28477	84.4878	69.05066	65.01564	65.36828	66.4782
F509W	86.06694	89.69982	83.58999	86.4523	67.67858	70.03068	66.17343	67.9609
F509L	86.3987	88.21282	83.89633	86.1693	67.67858	68.82748	66.98848	67.8315
F509S	85.40723	88.30491	85.44495	86.3857	66.33375	70.07084	67.39977	67.9348
F509Y	86.3987	87.68822	85.82213	86.6364	65.22731	65.67911	68.64884	66.5184
W511A	86.3987	84.46567	85.75809	85.5408	64.67858	65.32076	66.54944	65.5163
Y514A	85.73645	82.51252	86.07237	84.7738	67.31151	65.67707	66.78228	66.5903
extra loop	84.75258	81.02807	87.34107	84.3739	65.89144	63.48475	66.42312	65.2664

Table 31 Molecular weight analysis of SDS PAGE gel with recombinant toxins.
Image Lab quantification of molecular weight of 150 µg/mL (or 0.197 µM/ml) of activated Cry41Aa and recombinant toxins

6.6.2. Dose response of HepG2 cells treated with Cry41Aa recombinant toxins

A dose response analysis was carried out to assess the effects on cell viability after incubation with Cry41Aa and its recombinant mutants. The protein concentration of Cry41Aa and its mutants was optimised and visualised on SDS PAGE gel (figure 73) The concentration of the toxins in the medium of the 96 well plate was as follows 0.197, 0.131, 0.0657, 0.0263, 0,013, 0.006, 0.002 µM/ml.

Figure 74 is the graphical representation of the dose response CellTiter-Blue assay of HepG2 cell line incubated for 24 h with Cry41Aa as well as loop recombinant toxins. The following recombinant toxins F509A, F509L, F509S, F509W, F509Y, W511A, and the extra loop did not exhibit any cytotoxicity to HepG2 cells. This is based on CellTiter-Blue fluorescent assay where percentage cell viability is 100% relative to buffer regardless of toxin concentration. For statistical comparison of treatments in dose response experiments a Post-Hoc analysis applying a Bonferroni adjustment was carried out in SPSS.

Only Cry41Aa and the recombinant Y514A appear to be cytotoxic to HepG2 cell line. Here the fluorescent signal, indicative of the percentage of viable cells compared to that of buffer, was significantly less. In cells treated with the highest Y514A and Cry41Aa toxin concentrations (0.197 µM/ml) a significant reduction was noted in percentage of viable cells. Treatments with toxin concentration 0.197, 0.131, 0.0657,

0.0263 µM/ml of Cry41Aa and Y514A were statistically significantly when compared to treatments with buffer only cells. Furthermore, microscope observations confirmed that swelling of cells was observed within the first 10 min of incubation with Cry41Aa and Y514A.

Figure 75 illustrated a dose response curve for HepG2 cells treated with Cry41Aa and Y514A with the same concentration range (0.197 to 0.00263 µM/ml) for 24 h in order to estimate the half maximal lethal concentration (LC50) of both toxins. The LC50 values were determined by probit analysis using SPSS software. Cry41Aa had an LC50 value of 0.061 µM/ml, and Y514A had an LC50 value of 0.048 µM/ml. Y514A appears to kill faster with clear indication of cell swelling within 10 min of exposure to toxin mutant. The Post-Hoc analysis supports finding that the lower cell viability observed with recombinant Y514A treatment is significantly less when compared to treatments with Cry41Aa. This supports the theory that Y514A is likely more toxic than its wildtype Cry41Aa toxin.
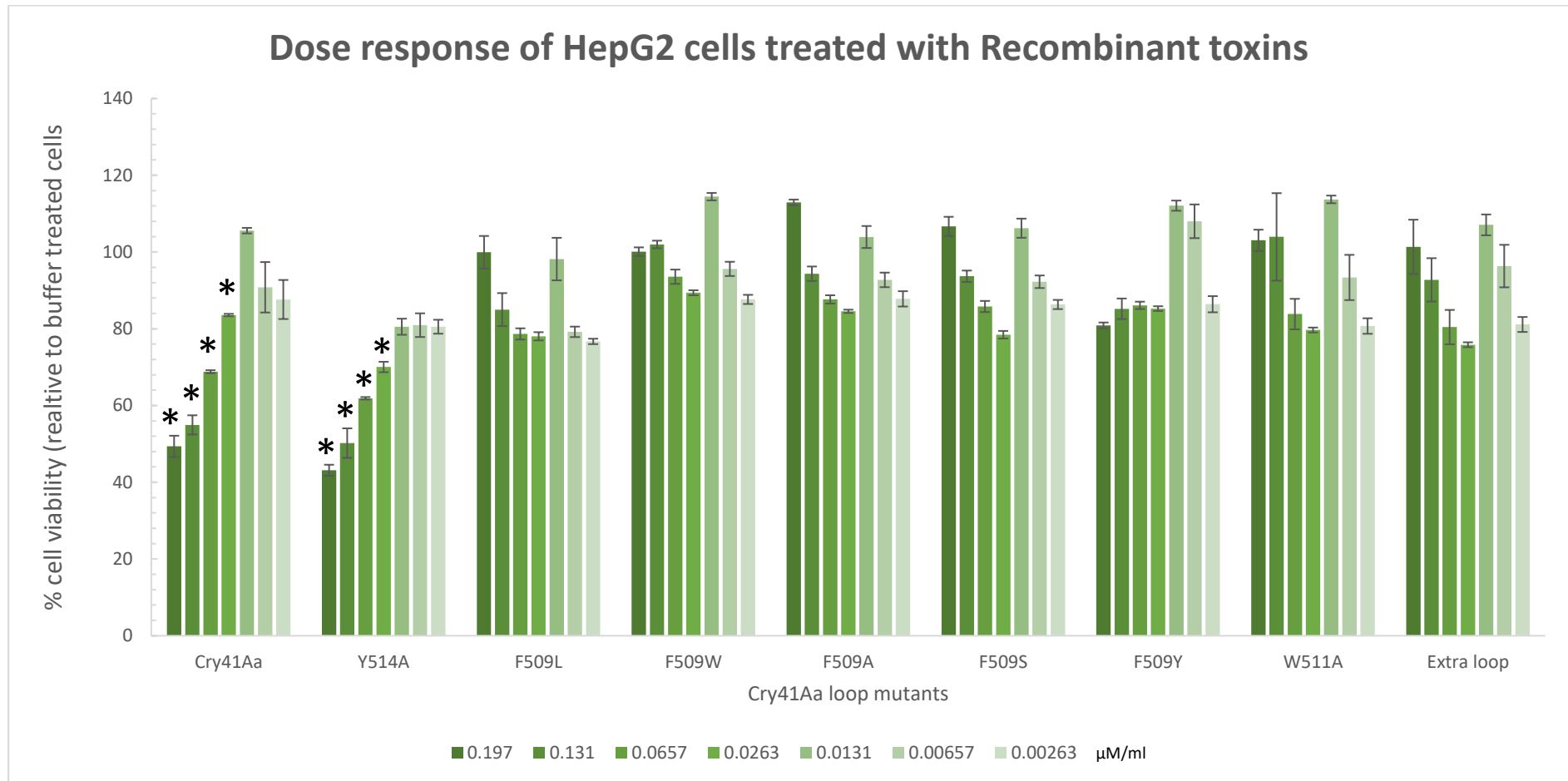
Figure 74 Dose response effects of recombinant toxins on HepG2 cells using CellTiter-Blue assay.
The cells were seeded 24 h before the day of experiment at 25x10⁴ cells/mL. Cells were treated with activated Cry41Aa, loop 3 and extra loop recombinants of different concentrations. Viability was measured as a metabolic activity of viable cells. Reading was taken 24 h post treatment with recombinant proteins. Bars with asterisks are significantly different from buffer only control cells (* p < 0.005, Post-Hoc comparison with Bonferroni correction).
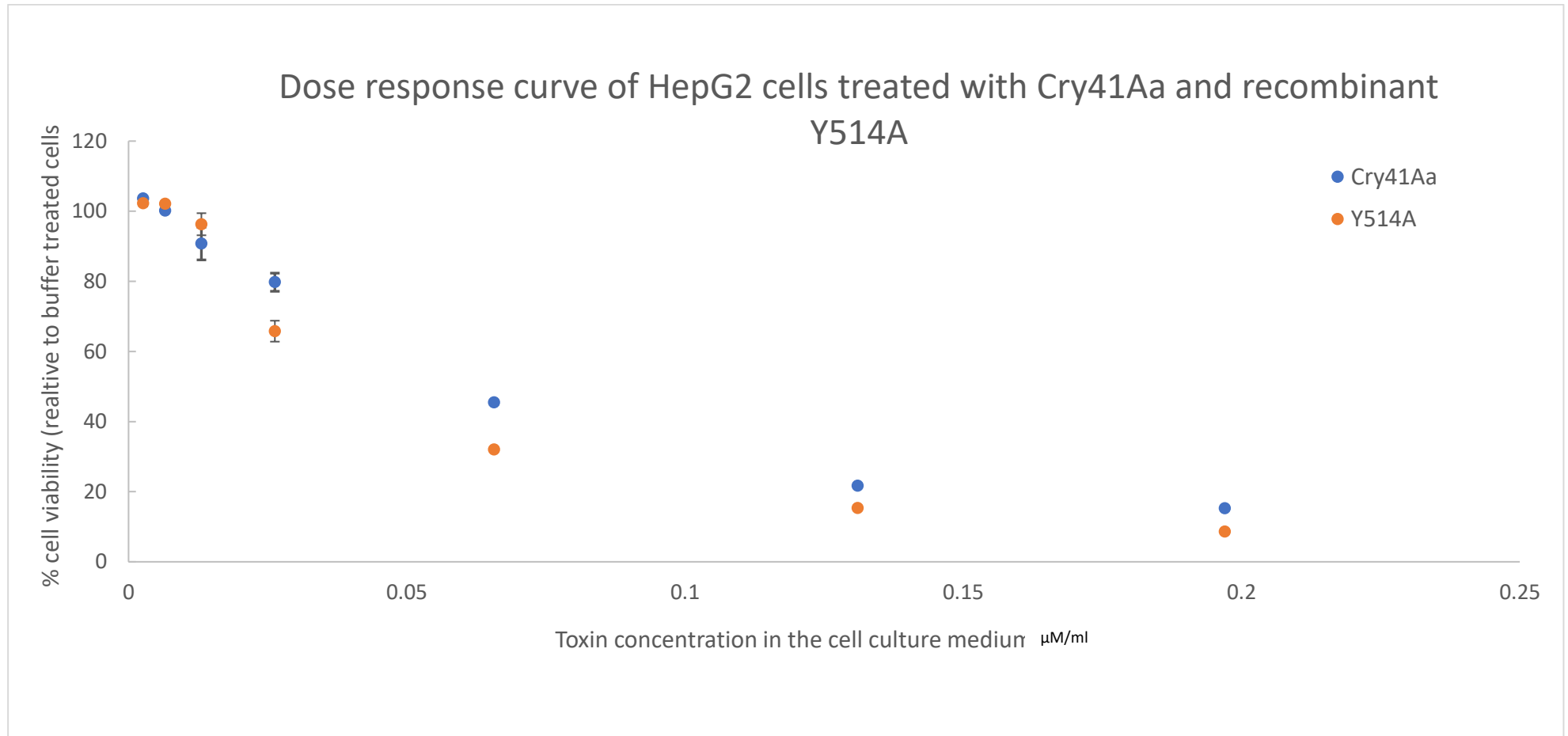
Figure 75 Dose response curve for HepG2 cells treated with Cry41Aa and recombinant Y514A.
 HepG2 cells were seeded at the density of 25x10^4 cells/mL. The following day cells were incubated with different concentrations of Cry41Aa and the recombinant toxin Y514A. Fluorescence was measured after 24 h using CellTitre-Blue. LC$_{50}$ was calculated using SPSS probit software Cry41Aa EL$_{50}$= 0.061 µM/ml, Y514A LC$_{50}$= 0.048 µM/ml.

The W511Y mutant did not produce a protease resistant core, however activated recombinant W511F did form a protease resistant core. The mutant W511F was engineered to investigate the effects of phenylalanine residue on toxicity. The protein concentration of this mutant was not optimised at the same time as the other degenerate residue substitution loop mutants, its concentration was calculated using Bradford protein assay. The effects of W511F on HepG2cell viability percentage were presented in graph as histograms in figure 76 after a 24 h incubation with W511F and relative to cell viability percentage in buffer. Recombinant toxin W511F was toxic to HepG2 cells. However, it is the protein concentration was not optimised with the cytotoxic Y514A and Cry41Aa on the same protein SDS PAGE gel.
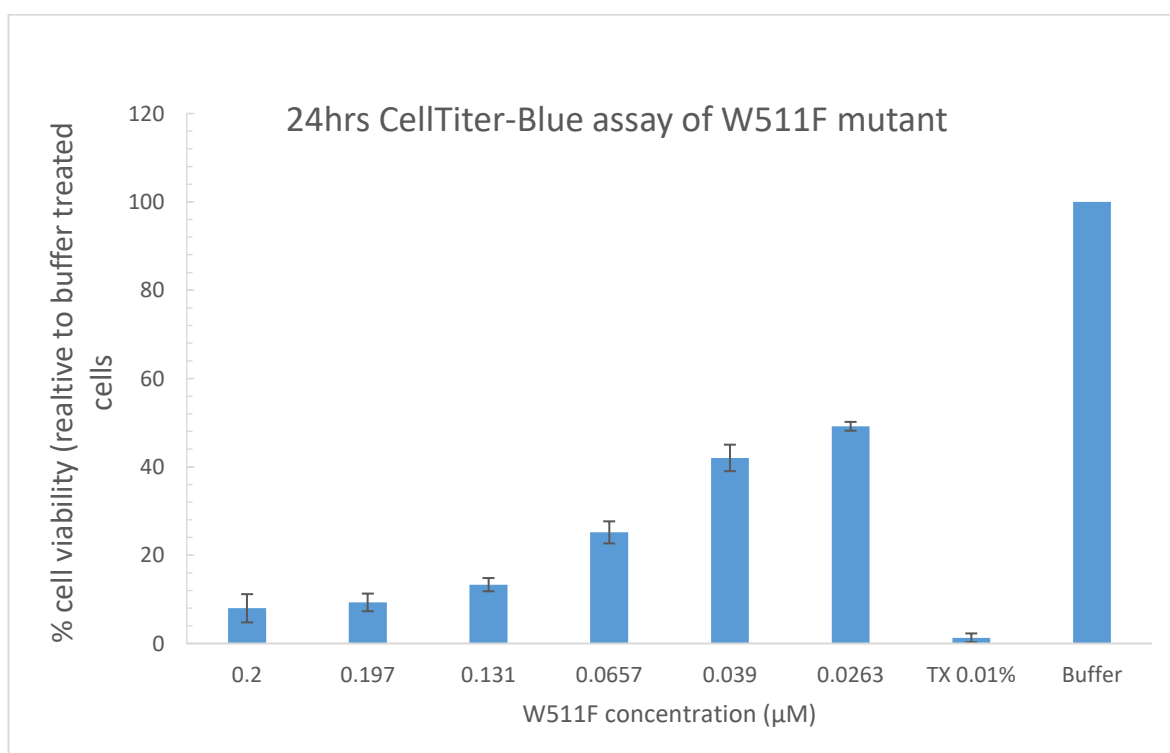
Figure 76 CellTiter blue assay of loop 3 recombinants
24hr CellTiter-blue assay of loop 3 substitutions W511F on HepG2 (seeded at 5,000 and 22x10^4). TX-100 act as positive control.

6.6.3. Western blot application to investigate p38 MAP kinase phosphorylation in HepG2 cell lines after incubation with mutant F509A

The loop 3 recombinant F509A has not shown any cytotoxic characteristics towards HepG2 cells. The reason behind this lack of toxicity by this single alanine substitute is unknown. The substitution may have induced loss interactions with HepG2 cell. It also possible that interaction continues to take place, but it no longer exerts the same effect, and thus does not that result in cell death. At the time of ligand blot experiments F509A was the only nontoxic mutant to have undergone two assays to assess its effects on HepG2 cells. CellTiter-Blue assay concluded that HepG2 cells were not affected after treatment with F509A. CellTox green assay indicated that nuclear membrane of HepG2 cells remained intact after F509A treatment and there was no different to buffer only treated cells.

HepG2 cellular response to F509A was investigated using ligand blotting and electrophysiology. It was carried out in an attempt to better understand the lack of toxicity of F509A and thus gaining a better understanding of the specificity of Cry41Aa. A cellular response by a target cell as a result of pore formation toxin can vary but one the commonly employed mechanism is the phosphorylation of p38 mitogen- activated protein kinase (MAPK) (Ratner *et al*., 2006). The MAPK protein family is made up of highly conserved serine-theronine kinases that are activated in response to extracellular stresses such as osmotic shock or DNA damage. One of the subgroups of the MAPK protein family is p38. When activated, p38 phosphorylates a number of proteins that

include cytokines, nuclear transcription factors and extra cellular receptors (Zarubin and Han, 2005).

A number of Cry toxins have been found to activate p38 pathways. These include p38 pathways induced by Cry5B toxin in *C.elegans*, as well as *M. sexta* and *A.aegypti* treated with Cry1Ab and Cry11Aa toxins (Cancino-Rodezno *et al*., 2010; Huffman *et al*., 2004; Porta *et al*., 2011). Domanska, confirmed that Cry41Aa induces p38 MAP kinase phosphorylation in HepG2 cells (Domanska, 2016).

Western blotting was carried out to examine if non-cytotoxic loop 3 recombinant F509A induced phosphorylation of p38 in HepG2 cell. Figure 77 shows a western blot prepared from HepG2 cell extract incubated with the different solutions. Sodium arsenite is a toxin known to induce p38 phosphorylation and is used here as a positive control, Cry1Ca an insecticidal 3-domain toxin is used as a negative control like the buffer. Cry41Aa is toxic to HepG2 (Krishnan *et al*., 2017), and loop 3 F509A mutant toxin which has so far not shown any signs of toxicity towards HepG2. CD59 is a loading control. It suggested that sodium arsenite, Cry41Aa and F509A all produced a signal for p38 MAP phosphorylation.

 A positive signal for F509A is unexpected as cell viability from CellTiter-Blue assay and membrane damage by CellTox-Green cell assay concluded that F509A is not toxic to

HepG2 cell lines. It is possible that F509A interacts with HepG2 cells causing a signal, but it also possible that this interaction is insufficient or unstable and does not result in cell death. Krishnan *et al*. (2017) had suggested that Cry41Aa has a similar mode of action to insecticidal 3 domain Cry toxins, and thus may cause pore formation in the membranes of HepG2 cells causing a p38 signal activation (Krishnan *et al*., 2017). However, it is the stability, size, and the time period that these pores remain open that may dictate if the cells die or recover. Western blots of HeLa cells and HL60 incubated with the F509A and Cry41Aa were carried out. Western blots of Hela cells did not detect any p38 MAP kinase signal (data not shown). However western blots of HL60 cell extracts indicated that both Cry41Aa and its mutant F509A did induce p38 MAP kinase phosphorylation as shown in figure 78. This finding is not reflective of cell viability data which indicated that both Cry41Aa and its mutants were not toxic to HL60 cell lines.
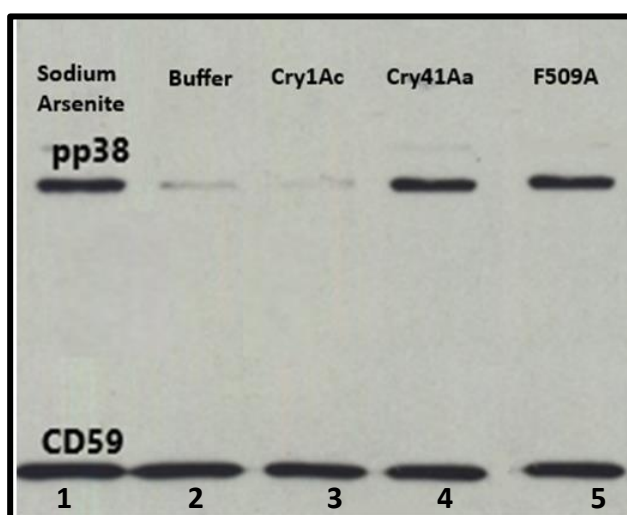

Figure 77 western blot analysis HepG2 cell extracts
Western blot of HepG2 cell extract treated with : Lane 1 sodium arsenite (0.5mM) lane 2 buffer, lane 3 Cry1Ac (15μg/mL), lane 4 F509A (15μg/mL) , and lane 5 Cry41Aa (15μg/mL), cells were lysed in NP-40 15 min after toxin treartment. 10μg of protein were loaded in each lane and after SDS-PAGE proteins were subjected to western blot analysis fro the detection of phosophorlayted p38 and CD59 (loading control).
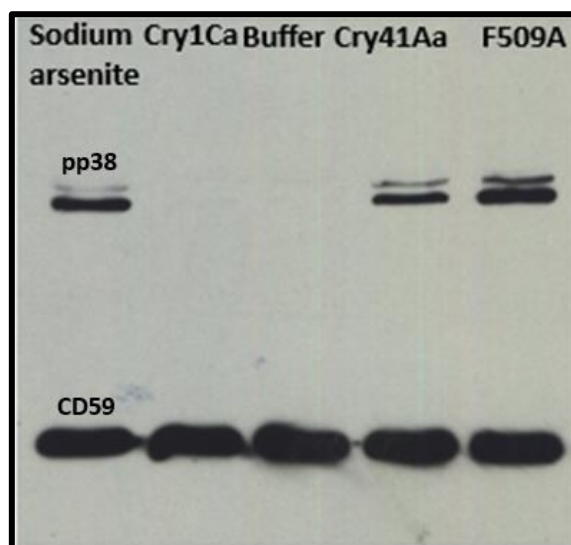
Figure 78 western blot analysis of HL60 cell extract
western blot of HL60 cell extract treated with: Lane 1 sodium arsenite (0.5mM). lane 2 Cry1Ac (15µg/mL). lane 3 buffer, lane 4 Cry41Aa (15µg/mL), , and lane 5 F509A (15µg/mL), cells were lysed in NP-40 15 min after toxin treartment. 10µg of protein were loaded in each lane and after SDS-PAGE proteins were subjected to western blot analysis fro the detection of phosophorlayted p38 and CD59 (loading control).

6.6.4. Electrophysiology application to investigate the interactions between HepG2 cell lines and the loop 3 recombinant F509A

Electrophysiology was employed to study the formation and type and of pores in membranes. The following section explores the effects of Cry41Aa and F509A on whole cell patch clamping and on an artificial planar lipid bilayer (PLB) membrane. Whole cell patch clamping technique uses cells whilst PBL uses artificial membranes. The following cell assay work was carried out by Barbra Domanska under the supervision of Prof Michelle West, university of Sussex, UK. Electrophysiology work was carried out by Barbra Domanska and Eva Fortea under the supervision of Prof Jean-Louis Schwartz, university of Montreal,Canada.

6.6.5. Whole cell patch clamping

Whole patch clamping is techniques discovered in 1976 used to study electrophysiology of membranes and ion channels (Neher and Sakmann, 1976). The nature of the ion channel formed in excised membrane patches or intact membranes is discoverable when recordings of the current passing through the membrane are made. In essence, ion channels are pores defined by their ability to conduct a current. The easier a current pass through a pore the more efficient the pore or ion channel is. In order to successful apply the patch clamp technique the membrane must first become electrically isolated from the external buffer by placing a glass pipette with electrolyte solution onto the membrane. An electrode attached inside the pipette records current flow in the membrane, the signal is picked up by an amplifier and data recorded.

Whole cell patch collects data on multiple channels on a given cell membrane. Here, the pipette sucks the intact cell membrane rapturing it but without disrupting contact between the outside of the cell membrane and rim of the pipette creating a seal. Eventually, the buffer in the pipette equilibrates with the cell cytoplasm. Currents data on whole patch carried out on whole cells tend to be big due the abundant number of channels opening simultaneously. Although whole patch clamping can record macroscopic currents in whole cell membrane, it cannot give detail on the nature and type of pore present.

Whole cell patch clamping was used to study macroscopic currents in HepG2 cells during incubation with Cry41Aa and loop 3 recombinant F509A. Data was collected prior to toxin addition and every 5 min thereafter for 20 min post toxin addition. The resting membrane potential of hepatocytes ranges between -30 to -40mV as a consequence of high basal chloride conductance (Moule and McGivan, 1990). KCl buffer, which consisted of 140mM KCl, 1.1 mM MgCl2, 0.1mM EGTA, and 10mM HEPES, pH 7.4, was added to the pipette and a holding potential of -20mV was applied so it was close to the resting potential of HepG2 cell membrane. NaCl buffer was added to the bath as the extracellular solution. A set of 17 one second depolarising step potentials (from -20 to +140mV from a holding potential of -20mV) were applied to initiate currents.

Figure 79 and 80 are graph representations of whole patch cell clamping on HepG2 cell membranes after exposure to cytotoxic Cry41Aa (figure 79) and its non-cytotoxic loop 3 recombinant F509A respectively (figure 80). The experiment aimed to measure the current that is exerted as a result of ion channels or pores opening as a consequence of toxin activity inHepG2 membrane. Three whole cell patch experiments were carried out on each toxin in order to obtain a mean value as shown in figures 73 and 74. It is clear that as the voltage increased current also increased in Cry41Aa treated HepG2 cells (figure 79). Significant current data is recorded for 15- and 20-min time points, however current increase as early as 10 min after toxin addition.

Figure 80 shows the effect of an electrical current on HepG2 cell membrane after exposure to F509A. It is clear that there is no elevated current recorded as a consequence of F509A toxin despite increased voltage and exposure time. The slight rise in current post a voltage of 70mV is due to the increase in voltage. This suggests that either F509A does not induce the significant opening of channels or that the effect is short lived rendering it unstable to have any significant effect on the membrane. This is consistent with permeability assay for F509A (cell Tox green, Figure 66 chapter 6.6.5) where results indicated that F509A had caused little damage to HepG2 cell membrane. This suggests inhibition of binding or ineffective pore formation by F509A.
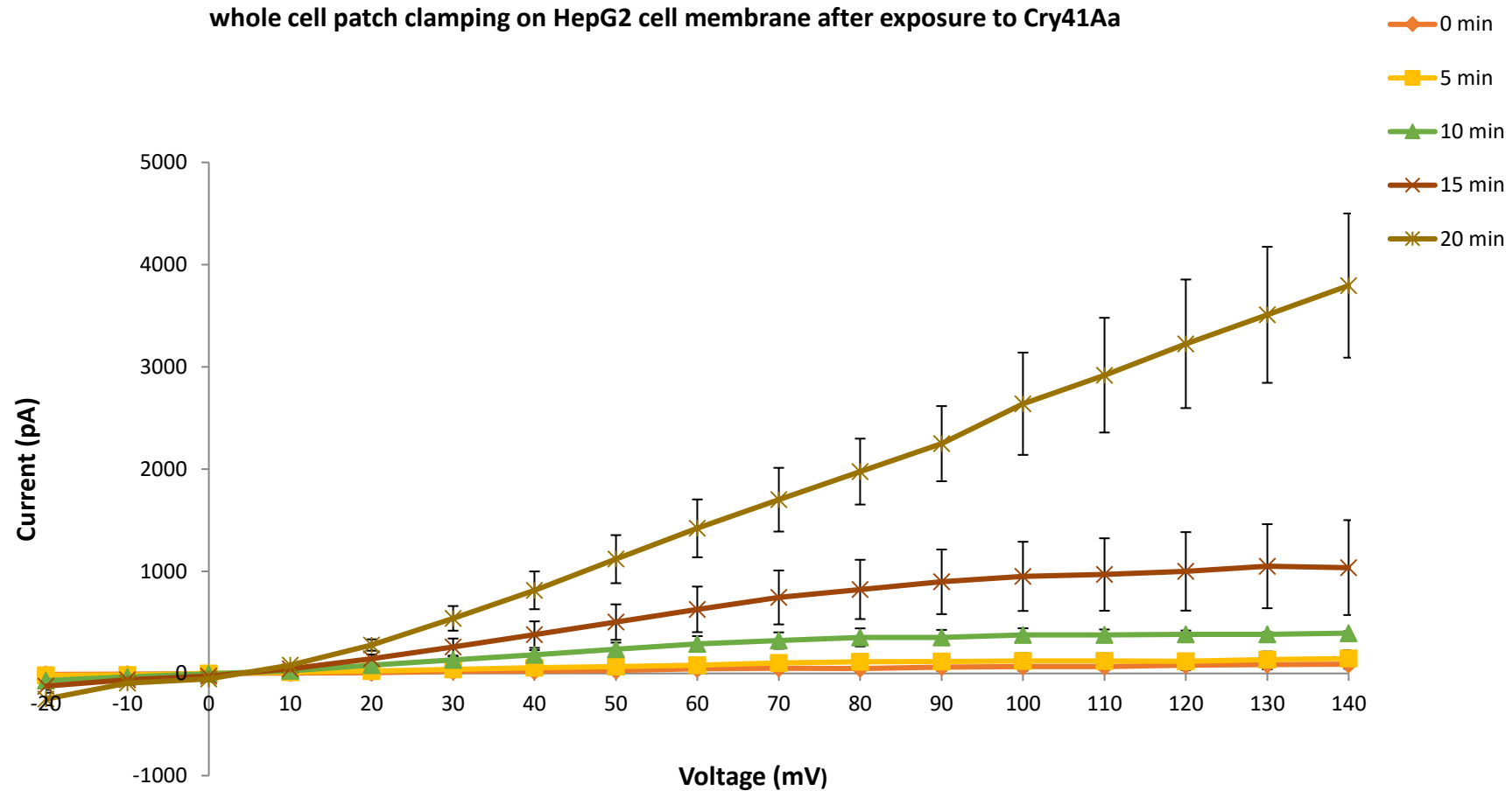
Figure 79 whole cell patch recording from a time course experiment of HepG2 cell exposed to Cry41Aa.
HepG2 cells were seeded at the density of 5x10^4 cells/mL on the glass coverslip inside the 35mm petri dish. The following day whole cell patch clamp data was recorded at 0, 5, 10, 15, and 20 min after exposure with Cry41Aa (12µg/mL) as well as NaCl bath solution. Currents were induced by 1 s lasting set of 17 depolarising potentials from -20 to 140mV form holding potential of -20mV. Error bars indicate the standard error of the mean. The lines represent the mean currents from three different whole patch cell experiments
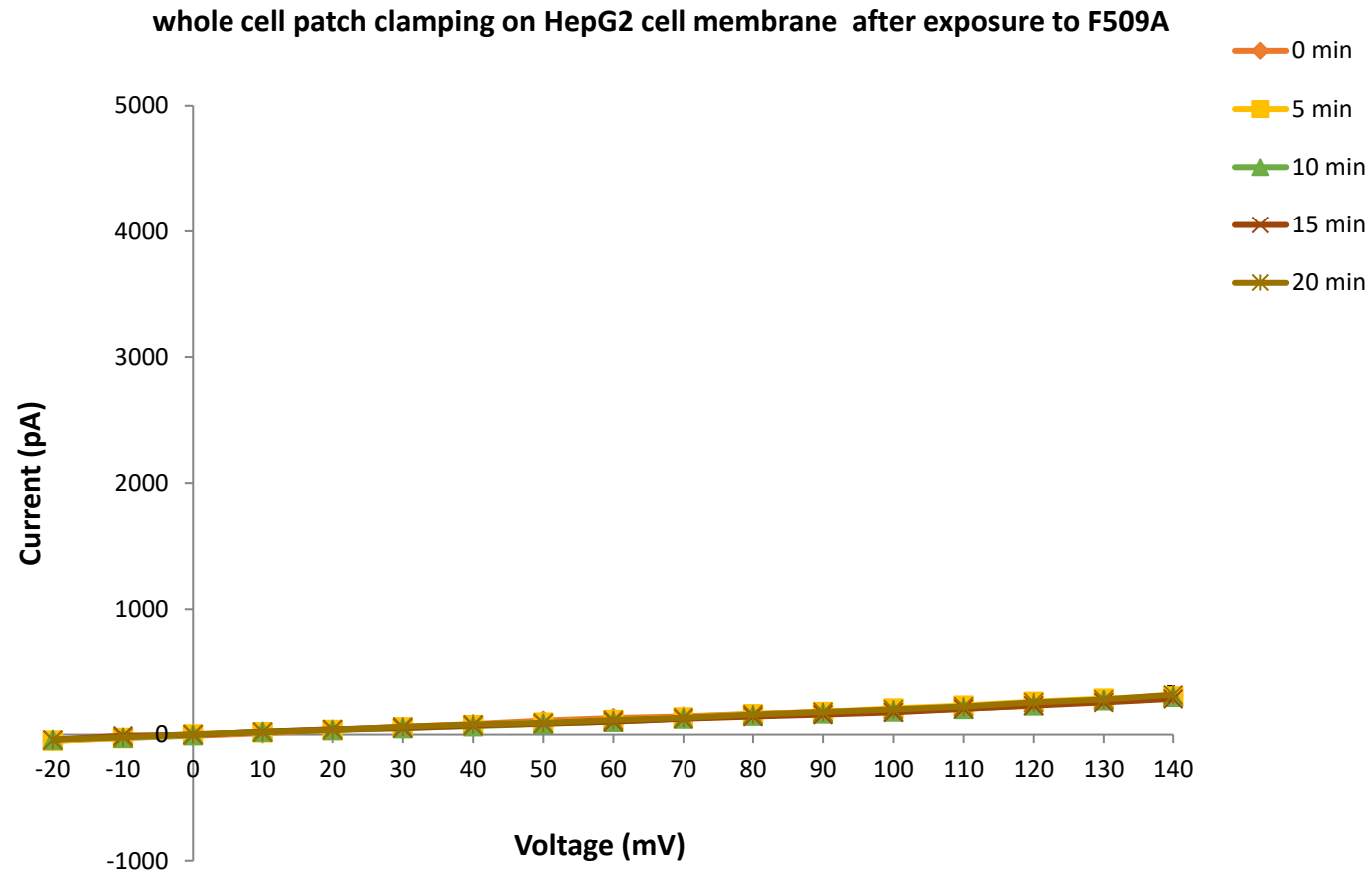
whole cell patch clamping on HepG2 cell membrane after exposure to F509A

Figure 80 whole cell patch recording from a time course experiment of HepG2 cell exposed to F509A.
HepG2 cells were seeded at the density of 5x10⁴ cells/mL on the glass coverslip inside the 35mm petri dish. The following day whole cell patch clamp data was recorded at 0, 5, 10, 15, and 20 min after exposure with F509A (12μg/mL) as well as NaCl bath solution. Currents were induced by 1 s set of 17 depolarising potentials from -20 to 140mV form holding potential of -20mV. Error bars indicate the standard error of the mean. The lines represent the mean currents from three different whole patch cell experiment

6.6.6. Single channel activity in planar lipid bilayers

Single channel activity can be investigated using an artificial planar lipid bilayer (PLB). An artificial phospholipids bilayer is prepared by using two identical aqueous solutions which represent the inner (inside the cell) and outer (outside the cell) regions of a membrane. This technique allows investigations to take place where the buffer and lipid arrangements can easily be manipulated due to the simplicity of this cell membrane model (Mueller *et al*., 1963). In vertebrate cells, the membrane is made up of phosphatidylcholine (PC) which is present in the external side of the membrane, phosphatidylethanolamine or PE which is present in the internal cytoplasmic side of the membrane and cholesterol adds to the fluidity of the membrane.

The experiment was set up as follows, two chambers known as cis and trans were separated by a lipid membrane which was pre-painted onto a plastic divider placed between the chamber. 1mL of buffer (150mM KCl, 1mM $CaCl_2$, 10mM HEPES, pH 7.5) was equally distributed between both cis and trans chambers. Traces of any current were checked for 30min before any of the toxins were added to ensure absence of channel activity. The toxins Cry41Aa or F509A were added to the cis chamber and activity was recorded at different applied voltages.

Three different experiments were carried out on both Cry41Aa and F509A. Electrical conductance is the measure of the ease at which a current pass through and this

measure was calculated to indicate the presence and effectiveness of pores. Figure 81 illustrates examples of traces recorded in planar lipid bilayer after application of Cry41Aa and F509A at a positive voltage. Any trace above zero indicated the opening of a channel. There is an obvious steady trace for Cry41Aa indicating that pores are structurally intact and functionally efficient. This is in contrast to F509A trace which remains close to zero and is not consistent suggesting that pores here were fragile and unstable.

Planar Lipid Bilayers Trace

Cry41Aa **(8µg/ml) at +100mV**

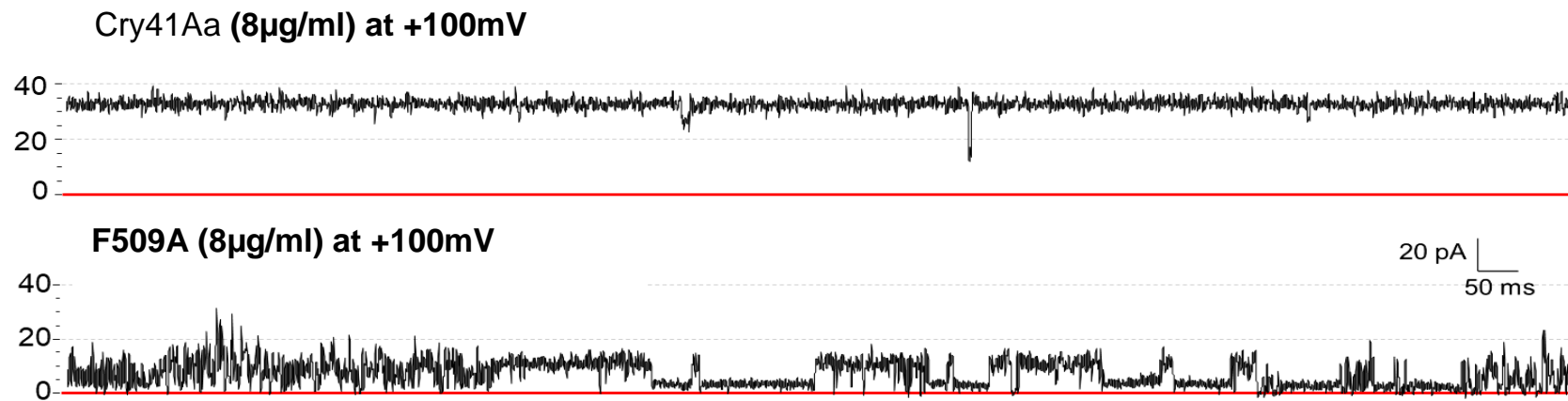

F509A (8µg/ml) at +100mV

Figure 81 sample of a single channel trace recoding in planar lipid bilayer after toxin application at the positive voltage.
 Single channel activity was recorded at +100mV (upper and lower traces) from PLB in KCl buffer after addition of Cry41Aa or F509A to the cis chamber. Records were filtered at 200Hz. The trace is representative of the three experiments

Approximately 20-25 current jumps were averaged for each experiment and analogous conductance values were grouped together. Data were plotted on a graph and a mean conductance within the group was read from a slope of linear regression.

Figure 82 is a graphical representation of the three experiments for Cry41Aa and the calculated mean values deduced from the linear regression points on the graphs and were carried out to investigate single channel activity as a consequence of Cry41Aa toxin. The data is consistent throughout the three experiments. A number of conductance values were recorded.

There is an indication that the two lowest conductance values may relate to highest conductance value. The two lowest conductance values roughly add up to the highest value. For example, in experiment I, the two lowest conductance values are 91 and 131pS which add up to the highest recorded conductance value of 220pS. It is possible that the lowest conductance values are due to two different channel populations that open and close independently, but when they are opened simultaneously, the highest conductance value is recorded.

Figure 83 is a graphical representation of the results of three experiments and the calculated mean values deduced from the linear regression points on the graphs carried out to investigate single channel activity as a consequence of loop 3 recombinant F509A

toxin. The conductance here is more diverse that for Cry41Aa. There are four different conductance values for each experiment suggesting the presence of more than one channel population.

There are no obvious patterns linking low conductance values with high conductance values in the same experiment and experiment II has quite a high value of conductance of 470pS. This is almost double the conductance value of the experiment I and III and it is not clear whether the PLB in experiment II has twice as many pores. No rectification was observed for data collected on both Cry41Aa and F509A PLB experiments as conductance did not change with the application of different voltages.

To investigate the type of ions that pass through the pore in i.e. the selectivity of the pore, chemical conditions were changed in the cis chamber. The concentration of KCl was increased to 450mM to create a concentration gradient; current was measured after this change to calculate whether the pore is catatonic or ionic.

The experiment was carried out thrice for each toxin. Figure 84 illustrates the selective conductance of the pores formed by Cry41Aa and F509A. Goldman-Hodgkin Katz potential equation was used to calculated the selectivity which states that ion permeability across a membrane is linked to reversal potential (Vr)(Goldman, 1943). The reversal potentials of Cry41Aa and F509A are calculated in table A and B of figure

84; Vr represents the potential at which there is no current passing through a pore. Both Cry41Aa and F509A maintained similar pS range for channel conductance in asymmetrical in symmetrical chamber conditions. However, the reversal potential shifts to -10.97 mV (± 1 mV) for Cry41Aa and a similar shift of Vr is observed for F509A with -11.5 mV (± 0.43 mV). This shift indicates that pores formed by Cry41Aa and F509A in the artificial PLB membrane tend to be more selective to cations. A small P_{K^+}/P_{Cl^-} is also observed with both toxins where Cry41Aa has a P_{K^+}/P_{Cl^-} value of 2.47(± 0.23), and F509A has a P_{K^+}/P_{Cl^-} value of 2.86 (±1.92). The large standard of error calculated for F509A in its P_{K^+}/P_{Cl^-} value is due to the instability of the trace as a result of instability and fragility of pores formed by the loop 3 recombinant F509A.
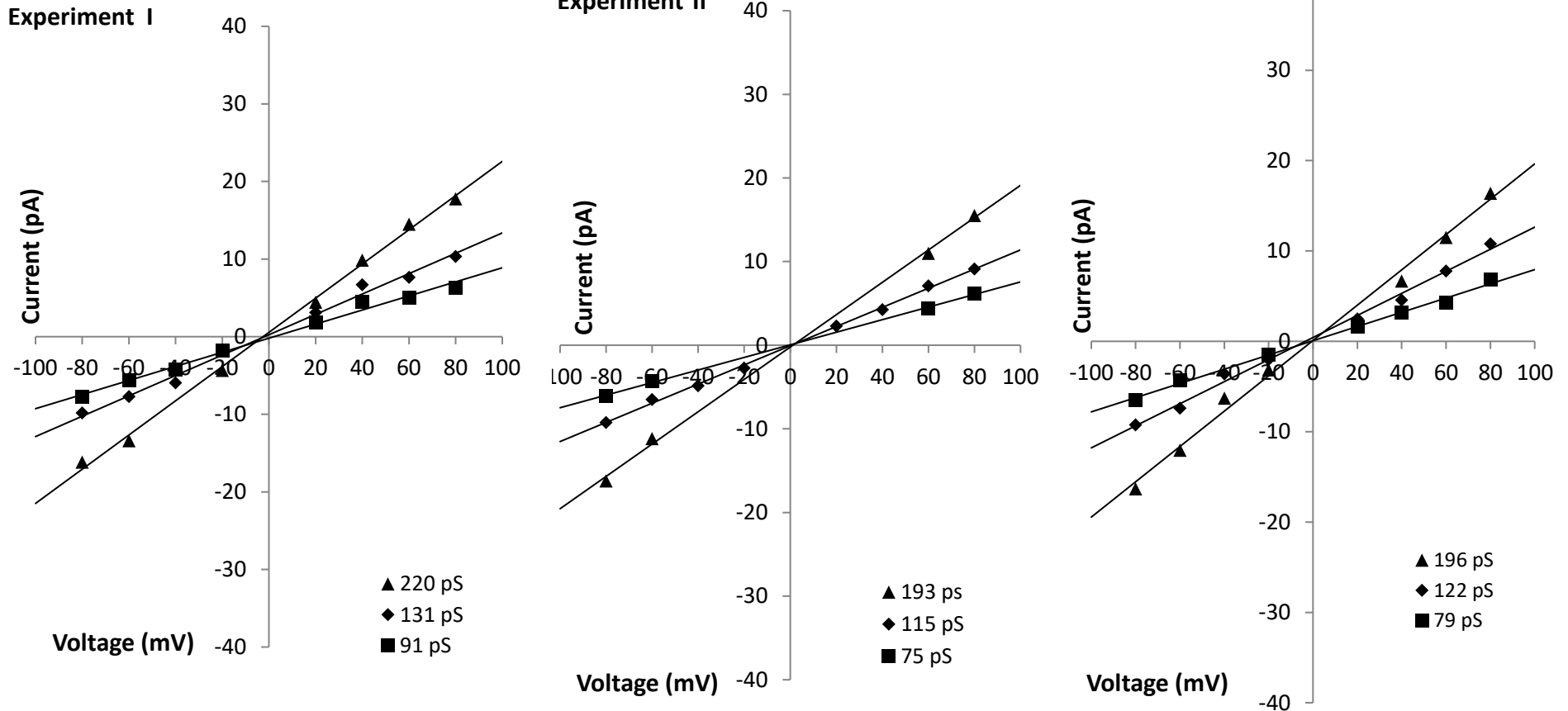
Figure 82 single channel –voltage relationship for Cry41Aa in planar lipid bilayer.

Current-voltage characteristics of single channel activity in 150/150 mM KCl (cis/trans) buffer (n=20-25) at different voltage after Cry41Aa (8μg/ml) was added to the cis chamber. Conductance was calculated (G=I/V) and data points fitted by linear regression. Experiment I: mean conductance values were deduced at the slopes of the linear regression on the data points as 91, 131, and 220pS. Experiment II: mean conductance values were deduced at the slopes of the linear regression on the data points as75, 115, and 193pS. Experiment III: mean conductance values were deduced at the slopes of the linear regression on the data points as 79, 122, and 196 pS. Data curtesy of Domanska, 2016

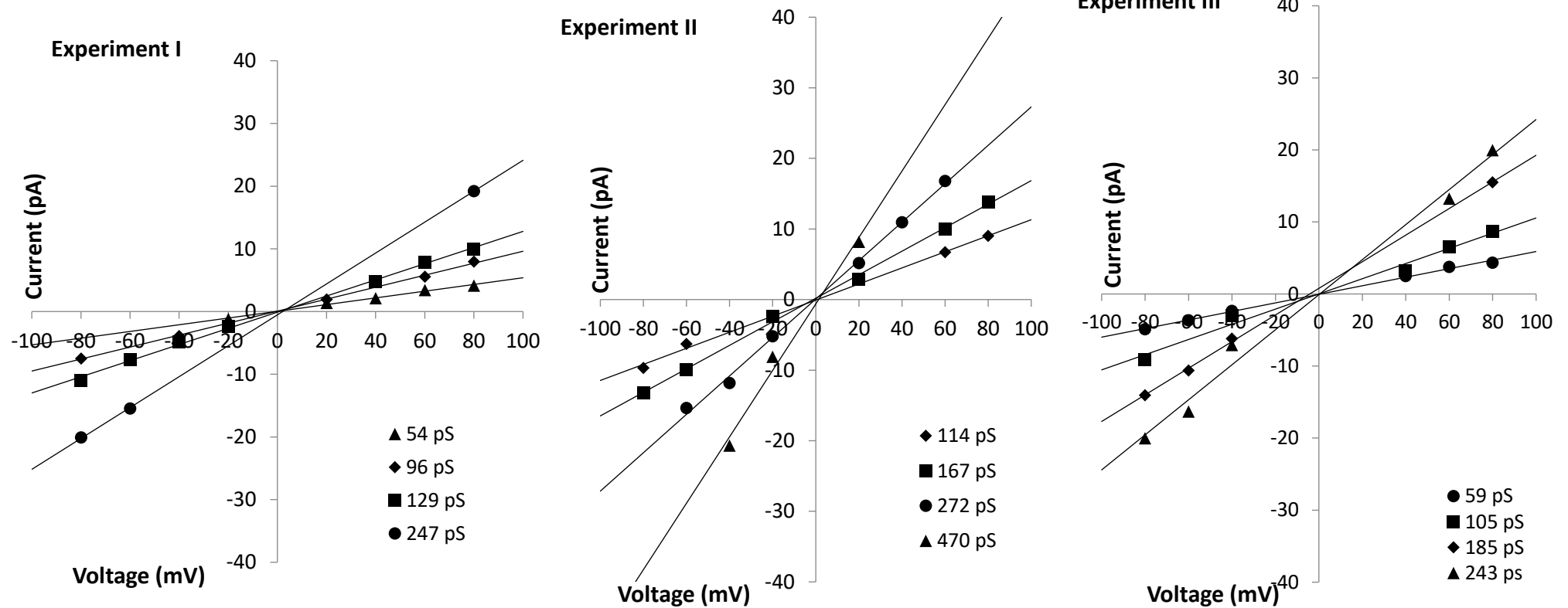**F509A: PLB single channel current activity**

Figure 83 single channel –voltage relationship forF509A in planar lipid bilayer.
Current-voltage characteristics of single channel activity in 150/150 mM KCl (cis/trans) buffer (n=20-25) at different voltage after F509A (8µg/ml) was added to the cis chamber. Conductance was calculated (G=I/V) and data points fitted by linear regression. Experiment I: mean conductance values were deduced at the slopes of the linear regression on the data points as 54, 96, 129, and 247pS. Experiment II: mean conductance values were deduced at the slopes of the linear regression on the data points as 114, 167, 272, and 470pS. Experiment III: mean conductance values were deduced at the slopes of the linear regression on the data points as 59, 105, 185, and 243pS. pS
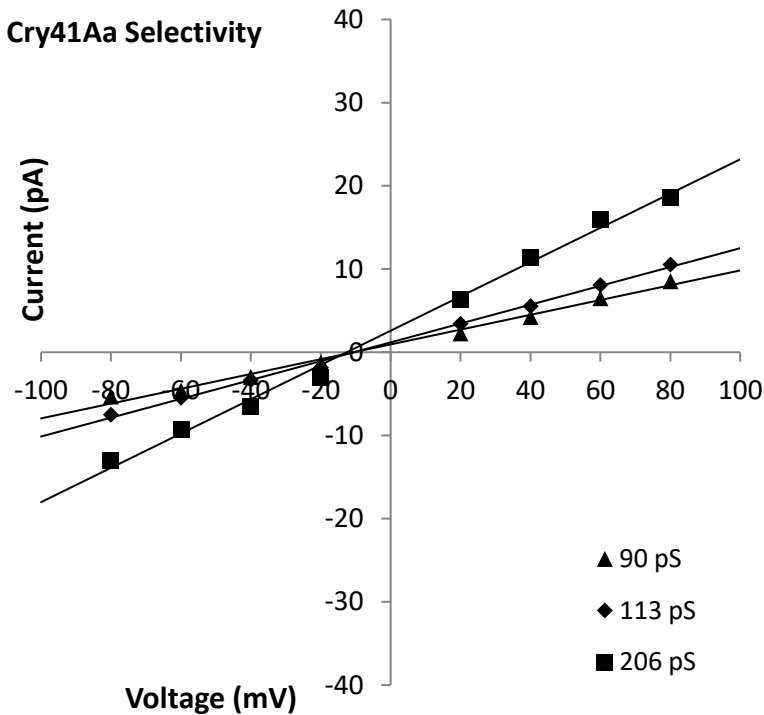
Selectivity of pores

A Experiment for Cry41Aa

	I	II	III	Mean	SEM
Vr (mV)	-10.5	-9.5	-12.9	10.97	1.0
PK^+/PCl^-	2.53	2.25	2.93	2.48	0.23

B Experiment for F509A

	I	II	III	Mean	SEM
Vr (mV)	-7.7	-14.0	-12.7	11.5	0.43
PK^+/PCl^-	1.85	3.30	2.88	2.68	1.92

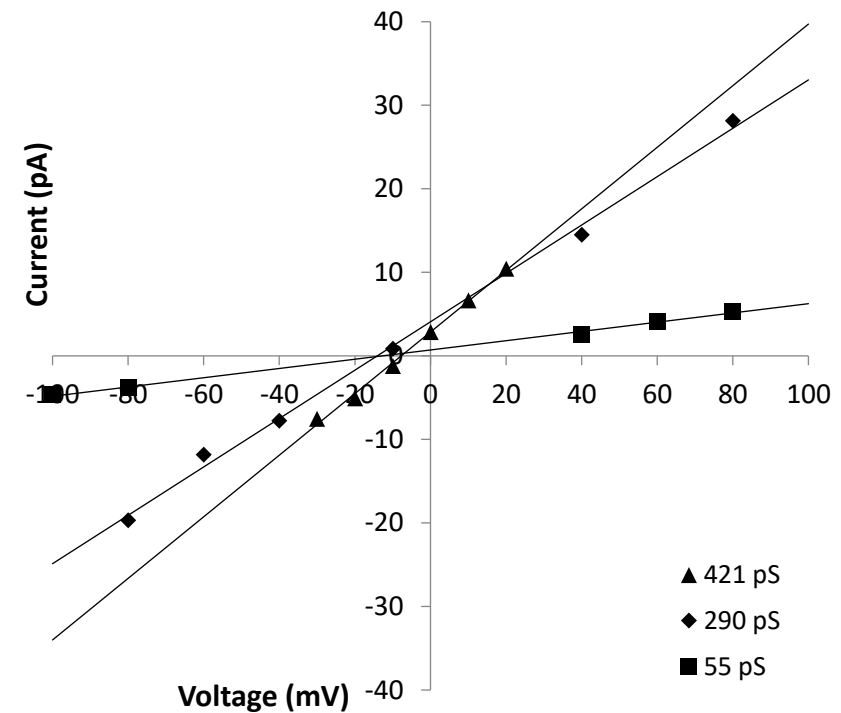**C   Cry41Aa Selectivity**

D   F509A Selectivity

Figure 84 single channel current –voltage activity for Cry41Aa **and F509A in planar bilayer asymmetric conditions for experiments I, II, and III.**
Selectivity (PK+/PCl-) and reversal potential (Vr) were calculated in PLB experiments I, II, and III for **(A)** Cry41Aa and F509A **(B)** in asymmetrical 450/150 mM KCl (cis/trans) conditions. Current-voltage characteristics of single channel activity in 450/150 mM KCl (cis/trans) buffer (n=20-25) at different voltage after toxin addition (F509A and Cry41Aa at 8µg/ml) was added to the cis chamber. Conductance was calculated (G=I/V) and data points fitted by linear regression. **(C)** Mean conductance values for Cry41Aa were deduced at the slopes of the linear regression on the data points as 90, 113, and 206pS. **(D)** Mean conductance values for F509Awere deduced at the slopes of the linear regression on the data points as 55, 290, and 421pS.

6.7 Discussion

Table 18 in chapter 6.2.2 summarises the constructs made via loop exchange, creation of hybrids with other Cry genes and loop residue substitution. The following Cry41Aa mutants were successfully solubilised and activated. Loop 3 exchange mutant pGEM1Acloop3 (this is made up of the Cry1Ac gene with loop 3 of Cry41Aa). The following are all loop substitutions, in loop1 the mutants ^{384}CVSC^{387}, ^{384}CLAC^{384}, and ^{384}GLAC^{384} were made. In Loop 3 substitutions the following mutants were made F509A, F509W, F509L, F509S, F509Y, W511A, W511F, W511Y, Y514A; and finally, in the extra loop substitution mutant ^{280}AAAA^{283} was created.

To explore the specificity of Cry41Aa and identify regions responsible a number of mutants were created as summarised in table 18. Several studies have applied site directed mutagenesis to investigate Cry toxin specificity and have implicated domain II loop in receptor binding and toxin specificity. In particular loops α8, loop1,2,and 3 of domain II are implicated in receptor binding and mutagenesis there can result in structural changes that affect the stability and susceptibility of the recombinant protein to protease such as trypsin (Dean *et al.*, 1996; Gómez *et al.*, 2003; Lee *et al.*, 1996, Schnepf *et al.*, 1998; Tuntitippawan *et al.*, 2005).

Likitvivatanavong *et al.* (2009) investigated specific residues in loops α8, loop1, 2, and 3 of Cry11Ba for binding. The alanine residue substitutions in loops α8, loop1, and 3 resulted in reduced toxicity by 80-100% towards *Aedes aegypti* and *Culex* fourth instar

larvae (Likitvivatanavong *et al* ., 2009). Specific residue in $\alpha 8$ of Cry1Ac were substituted for alanine resulted in a reduced binding affinity for BBMV and the purified receptor aminopeptidase N (APN) from both *Lymantria dispar* and *Manduca sexta* and subsequent reduction in toxicity (Lee *et al.*, 2001).

Further studies on binding and toxicity were carried out which involved mutagenesis. The Cry1Ab residues of loop 3 (toxic *to M. sexta and H. virescens*) underwent alanine residue substitutions. The resulting mutants were stable in trypsin and insect gut juices and were thought to be structurally similar to wildtype. However, bioassay data revealed significant reduced toxicity and binding analysis of BBMV from both insect's species. This loss and reduction of toxicity in some mutants may be due to the reduced initial binding.

The study highlighted the role of hydrophobic residues where it was observed that alanine substitutions in loop 3 of Cry1Aa did exhibit reduce toxicity as a consequence of reduced initial binding of mutant toxins to *Bombyx mori* and *M. sexta* BBMV. The study implicated hydrophobic residues in loop 3 of both Cry1Ab and Cry1Aa in initial binding to receptor molecules that proceed to form hydrophobic interactions (Rajamohan *et al.*, 1996b).

Mutants which successfully produced crystals were harvested and characterised against wildtype Cry41Aa. Schematic table 18 in lists all the mutants made from Cry41Aa including those which did not proceed to cell assay stages. The recombinants W511Y

and W511F were created and processed in the same manner as Cry41Aa and other loop recombinants. The recombinant W511Y was not stable in trypsin and no further action was taken. The W511F recombinant did result in two proteolytically stable proteins of ~80 KDa and ~60 KDa ; however, its protein concentration was not optimised against the control Cry41Aa. Despite this, its toxicity and that of loop1 recombinants were assessed (see chapter 7) with the remaining loop mutants.

Yamashita *et al*. (2005) solubilised the parasporal inclusions obtained from *Bt* strain A1462. A number of major peptides 180, 150, 120, and 100 KDa were observed; however, the most significant was the 88 KDa doublet. The study findings proposed that it was from this doublet that the 64 KDa cytotoxic peptide and the minor 80 KDa peptide were derived from protease treatment. They were able to successfully isolate each protein and observed that was indeed the 64 KDa peptide that is toxic to HepG2 cell line. In the present study Cry41Aa was solubilised in sodium carbonate and DTT. A protein of ~ 80-88 KDa which is thought to be encoded by ORF2 was consistently observed with better dissociation when solubilised in the presence of DTT. A protein of ~ 120 KDa was thought to be encoded by ORF3. Upon treatment with trypsin, the solubilised Cry41Aa degrades to a protein of ~ 80 KDa (or slightly less as) well as protein of ~ 60 KDa (Krishnan *et al*., 2017). ORF3 band was absent after trypsin activation. It is thought that this peptide lacks a protease resistant core and or is unstable in trypsin thus degrading completely upon toxin activation.

In this study, AKTA was applied in an attempt to isolate each activated protein after trypsin treatment. However, it was not possible to isolate the two peptides produced by each activated protein from Cry41Aa or any of its mutants. Thus, the main method of protein purification was through the use of gel resin filtration column which allowed purification of large volumes of activated protein.

Previous attempts to express Cry41Aa without ORF3 resulted in an unstable protein that did not solubilise well and could not be activated by proteases (Krishnan, 2013). The ORF3 of Cry41Aa has similarities with the C-terminal part of the larger 3-domain Cry toxins, a region known to be responsible for expression and crystallisation of toxins in *Bt* (de Maagd *et al.*, 2003). This suggests that ORF3 gene may plays a role in the crystallisation and correct folding of the Cry41Aa crystals (Krishnan *et al.*, 2017).

All recombinants produced crystals which were indistinguishable from those made by Cry41Aa when observed under both light and electron microscopes. In addition, the recombinants are indistinguishable from the wildtype Cry41Aa on an SDS PAGE gel. They were processed similarly to the wildtype and no observations were made about any differences in their response to conditions that would differentiate them from wildtype Cry41Aa.

F509C and W511Y were the only Cry41Aa loop 3 recombinant which produced crystals but failed to solubilise well and were degraded upon activation by trypsin. These recombinants may have a different folding conformation to Cry41Aa. The stability of

recombinant toxins may possibly be affected by the type of residue at a particular position with consequently could increase sensitivity to trypsin.

The recombinant mutants of Cry41Aa had their concentration optimised in preparation for cell assay work. This stage of the study was addressed by bringing together analysis data from different tools to combine and support findings. The Bradford protein assay was able to provide a numerical value of the protein concentrations. SDS PAGE gels revealed visually these numerical representations and showed how accurate the values were to their numerical representation. The image lab software allowed for the quantification of toxin concentration from SDS PAGE gel bands densities relative to a control bands density as well as quantifying the average molecular weight (KDa) of recombinants against a protein marker of the same SDS PAGE gels.

From these data it was concluded that the recombinant toxins have a similar molecular weight to that of wildtype Cry41Aa. It has also confirmed that the stock 150 μg/mL concentration of each recombinant toxin was very close to the control band densities of Cry41Aa, where each of the reference bands have a value of 1 the recombinants show an average value of 0.9 to 0.8 when compared against the reference bands of three gels. This confirms that protein concentration of the stock samples for each recombinant were optimised and finalised as 150μg/mL, before proceeding to cell assay analysis.

Initial mutagenesis of Cry41Aa loop 3 highlighted the residues at F509 and W511 and their potential to influence the specificity of Cry41Aa. Cell assays were carried out on

the three initial loop 3 recombinants F509A, W511A, and Y514A. CellTiter Blue assays indicated that both F509A and W511A did not exhibit toxicity towards HepG2 cell lines. However, Y514A retained toxicity and showed signs that it killed faster and more cells within the same given time compared to wildtype Cry41Aa. A dose response on HepG2 cells treated with Cry41Aa and Y514A confirmed an EL_{50} for both toxins. Cry41Aa had an EL_{50} value of 0.061 µM/ml, and the recombinant Y514A had an EL_{50} value of 0.048 µM/ml. This confirmed that Y514A is indeed a more potent toxin than wildtype Cry41Aa.

The lack of toxicity by F509A was confirmed by cell assay with cytotoxic markers (Cell Tox Green) which binds nuclear material once the cell membrane has been damaged. Here, no significant cell membrane damage to HepG2 cells was recorded for mutant F509A. It was clear that alanine substitutions at position 509 had knocked out the toxicity, however it did not explain whether this loss of toxicity was due to loss of the wildtype residue as a result of the substitution or if the position of the residue formed an integral part of Cry41Aa's specificity. Furthermore, it was not known whether the interaction between F509A and HepG2 cell membrane did occur but that this interaction no longer produced a toxic effect on the cell. In the quest to better understand the specificity of Cry41Aa further amino acid substitutions were made at position 509 and 511 of loop 3.

None of the mutants created in position 509 were observed to have any toxicity to HepG2 cell lines. The mutant W511F was toxic to HepG2 cell lines.

The cytotoxicity of the initial three mutants F509A, W511A, and Y514a as well as wildtype Cry41Aa were tested on additional cell lines. BL31 showed only slight sensitivity to Y514A and Cry41Aa. Other cell lines did not show susceptibility to either Cry41Aa or Y514A.

Assays on HL-60 indicated that trypsin activated Cry41Aa and its recombinants did not have a toxic effect on this cell line. This finding contradicts previous HL-60 toxicity data collected on proteinase K activated Cry41Aa by Yamashita *et al.* (2005). Therefore, experiments on Cry41Aa focused on the use of HepG2 cell lines instead of HL-60. Ongoing research on the toxicity of Cry41Aa confirmed its toxicity towards HL-60 cell lines only when activated by proteinase K (Souissi, 2018). These conflicting findings put forwards the argument that specificity may be influenced by how toxins are activated.

Ongoing research on the toxicity of Cry41Aa confirmed its toxicity is observed towards HL-60 cell lines only when activated by proteinase K (Souissi, 2018). These conflicting findings put forwards the argument that specificity may be influenced by how toxins are activated. The cancerous Hela cell line was a negative control for Cry41Aa, however previous studies have shown that high dose of Cry41Aa can have a toxic effect on Hela cells (Yamashita *et al*., 2000). This postulation that Cry41Aa has the ability to from pores and may even bind non-specifically. Low selectivity values for both Cry41Aa and F509A channels from electrophysiological analysis supports this notion and provides support to the theory that Cry toxins did not evolve with or alongside mammalian cells.

Electrophysiological technique and ligand blotting were used to study the effects of Cry41Aa and F509A on HepG2 cell membrane. One of the responses by a target cell as a result of interactions with a pore forming toxin can be the phosphorylation of p38-mitogen activated protein kinase (Zarubin, T. and Han, J. 2005).

Western blots carried out on cell extracts after incubation with Cry41Aa and F509A indicated that both toxins induced a positive signal for p38 phosphorylation. Electrophysiological analysis indicated that Cry41Aa is a pore forming toxin which is potent at 0.5-1.05 µM/ml concentration for optimal pore formation effect. It forms very stable membrane pores with an opening probability of 1 (Domanska, 2016). Conductance of Cry41Aa range between 75-220 pS and the pore shows no rectification, but the pore seems to be voltage dependent and is slightly cationic selective (Rv: -10.96 mV).

Similarly, the loop 3 recombinant F509A is able to form pores in synthetic PLBs, however in contrast the pores are less stable and display different kinetics and opening probability. Conductance of F509A is in the range of 54- to 470. The selectivity of F509A is similar to Cry41Aa however standard error here is much larger (1.92) due to instability of the pore. PLB pores formed by F509A are also slightly cationic, a characteristic observed in Cry1B, Cry1Ac, and Cry3A (Schwartz *et al*., 1997; Slatin *et al*., 1990; Walters *et al*., 1993) .

In conclusion, F509A did not exhibit toxicity to HepG2 cell as evident by CellTiter blue data where cell viability percentage was no different to that of the buffer. In addition,

the membrane damage assays revealed that F509A did not cause significant cell membrane damage. These findings although supportive of F509A lack of toxicity to HepG2 cells did not explain the cause of this loss of toxicity.

Western blots carried out on F509A indicated that it did induce phosphorylation of p38. From the results, it was evidence that F509A did indeed bind or interact with HepG2 cells however the binding was ineffective or insufficient and did not result in cell death. Electrophysiological analysis supported this finding and confirmed that although F509A did form pores in the cell membrane these were unstable or were few in number.

It is important to note that membrane damage by F509A was enough to induce a p38 signal, similar to that observed with HL60 cells extract after incubated with F509A despite cell viability data indicating that F509A was not toxic to HL60 cell lines. Therefore, it is speculated that F509A binds to both HepG2 and HL60 cells. This binding leads to impaired pore formation in the cell membrane promoting phosphorylation of p38. It is possible that pores are repaired by cells or there are only few pores which do not induce toxicity, and thus not detectable in cell viability or membrane damage assays. F509A was one of loop 3 mutants created at position 509. All the mutants at this position did not exhibit toxicity to both HepG2 and HL60 cell lines. It is reasonable to suggest that like F509A, all the other mutants at this position may exhibit similar mode of action, by binding to cell in a manner that is ineffective and which results in unstable pore formation and thus lack of toxicity.

Phenylalanine was introduced at position W511 to explore whether the nature of the residue or its position could affect a toxin's ability to be potent. The recombinants W511F retained toxicity to HepG2 cells (figure 76 chapter 6.2.2). Further attempts to introduce more amino acid substitutions at position 511 resulted in mutant proteins that were proteolytically unstable. Phenylalanine is an aromatic hydrophobic residue; these characteristics have been implicated in the toxicity of Cry toxins (Pardo-Lopez *et al.*, 2009; Vachon *et al.*, 2012; Rajamohan *et al.*, 1996b; Dean. and Sylvis, 2006). However, substitution with any aromatic hydrophobic residues does not guarantee a proteolytic resistant mutant protein as was seen with mutant W511Y which was unstable in trypsin.

7.0 General finding and Discussion

7.1 Summary of findings

This chapter will summarise the findings from the previous chapters and discuss their significance. It will conclude by evaluating the methods employed and the future work that can be done to continue exploring Cry41Aa and exploit its benefits.

The 3-domain Cry toxins have been explored for their ability to kill certain insect orders. Their narrow killing range emphasised their specificity to target cells and *Bt* based pesticides have been widely used. The 3-domain toxin specificity has been investigated in the quest to understand the mode of action of Cry toxins, create novel toxicities, and overcome resistance (Crickmore *et al*., 1998; de Maagd *et al*., 2001; Bravo *et al*., 2011).

Mizuki *et al*. (1999) investigated the parasporal inclusions from *Bt* cells that appeared to be non-insecticidal and did not display any biological activity, however they did display very specific cytotoxicity to some cancer lines (Mizuki *et al*., 1999; Mizuki *et al*., 2000). This unique category of Cry toxins was named parasporins and constitutes , parasporins 1,2,3,4,5, and 6 (Okumura *et al*., 2010). Parasporin 3 or Cry41Aa was first reported by Yamashita *et al*. (2005) as a 3-domain Cry toxin with

cytotoxic properties that kills hepatic and cervical cancer cell lines (HepG2 and HL60). This very specific ability to identify and kill these two cancer cell lines raised the question about the mode of actin of Cry toxins.

Cry41Aa shares the five conserved blocks typically found in insecticidal 3-domain Cry toxins. Yamashita *et al.* (2005) demonstrated that it also required *in vitro* solubilisation and activation similarly to insecticidal 3-domain Cry toxins in order to obtain the protease resistant cores known to be toxic to the cancer cell lines. It has a ricin domain which resembles the β-trefoil domain found in putative carbohydrate binding epitopes, however Krishnan, (2013) confirmed the ricin domain was not responsible for Cry41Aa's specificity to HepG2 cell lines, as when it was removed, Cry41Aa retained cytotoxicity (Krishnan, 2013).

Many studies on the specificity of Cry toxins have highlighted the exposed loops of domain II (de Maagd *et al.*, 2001). Research findings show that it is domain II which is primarily responsible for receptor recognition and target cell specificity in 3-domain Cry toxins (Bravo *et al.*, 2007; Crickmore *et al.*, 1998; de Maagd *et al.*, 2001; de Maagd *et al.*, 2003).

Yamashita *et al.* (2005) *a*rgued that Cry41Aa may have a similar mode of action to a pore forming 3-domain insecticidal Cry toxin (Krishnan *et al.*, 2017; Yamashita *et al.*, 2005).

This comparison can be extended to question the extent of the similarity of Cry41Aa to insecticidal Cry toxins and how research on insecticidal 3-domain Cry toxins sheds light on the specificity of Cry41Aa.

This thesis aimed to investigate the specificity of Cry41Aa to HepG2 cells through mutagenesis in domain II loops. It employed bioinformatic tools to narrow down and identify regions in Cry41Aa that were potentially associated with binding and receptor recognition. Bioinformatic tools collated and compared data from insecticidal Cry proteins to predict sequences in domain II loops and model the secondary of Cry41Aa. The predicted domain II loops of Cry41Aa were then subjected to a number of mutagenesis approaches.

First, a loop 3 exchange loop was created with loop 3 of insecticidal Cry1Ac. The resulting recombinant protein was unstable and did not result in a protease resistant protein. The Cry1Ac insecticidal recombinant with loop 3 of Cry41Aa did result in a stable protease protein. Cell assay analysis indicated that this mutant did not gain toxicity to HepG2 cell. Second, a number of hybrids of Cry41Aa OFR 2 were created with insecticidal 3 domain Cry toxins.

Cry41Aa is a split toxin made up of ORF2 and ORF3. The five conserved blocks that make up domain I, II, and III and are typically found in 3-domain insecticidal. Cry were identified in ORF2 of Cry41Aa. A number of domain hybrids of Cry41Aa with insecticidal

Cry domains were created. Unfortunately, none of the domain hybrids resulted in protease resistant proteins.

Third, a number of substitutions in loop 1 and 3, the extra loop of Cry41Aa were created. In loop 1 a number of degenerate substitutions were created. This resulted in three protease resistant proteins which were assessed for their toxicity to HepG2 cells. The three mutants retained toxicity to HepG2 cells and were indistinguishable from wildtype Cry41Aa on an SDS PAGE gel.

In loop 3 of Cry41Aa three initial substitutions were made in position F509, W511, and Y514. These aromatic residues were substituted with an alanine residue. All three mutants resulted in protease resistant proteins and their toxicity to HepG2 was analysed. The Y514A retained toxicity and HepG2 cell assays suggested that is more potent than wildtype Cry41Aa. The remaining F509A and W511A lost toxicity to HepG2 cells. In an attempt to shed light on the specificity of Cry41Aa, the position of residues and type of residues were investigated further.

A number of degenerate residue substitutions were made at position 509. All degenerate substitution mutants resulted in a protease resistant protein with the exception of F509C. Two aromatic residue substitutions were created in position 511, W511F and W511Y. The W511Y mutant did not result in a protease resistant protein.

Substitutions in loop 3 resulted in 11 recombinant protease resistant proteins.   A multiple alanine substitution mutant was created in the extra loop of Cry41Aa.

12 recombinant protease resistant proteins, including the extra loop mutant, were analysed and compared to wildtype Cry41Aa. Transmission electron microscopy revealed that 12 mutants produced crystals which were undisguisable from those made by Cry41Aa. SDS PAGE gels of solubilised, and trypsin activated recombinants confirmed that each of the 12 mutants resulted in two protease resistant proteins of ˜60 and ˜80 KDa. These proteins were indistinguishable from those produced by activated wildtype Cry41Aa. Along with activated Cry41Aa, the 12 mutants were purified, and concentrations optimised in preparation for HepG2 cell assays.

All mutants did not exhibit cytotoxicity to HepG2 cell lines with the exception of W511F and Y514A. Membrane assay of the effect of F509A on HepG2 cells indicated that it did not damage cell membrane, however, western blots of the effects of F509A on HepG2 cell lines indicated the activation of p38 MAPK signalling pathway. This is known to be activated in response to pore formation (Huffman *et al.*, 2004).

Electrophysiology analysis revealed that both F509A and Cry41Aa form pores in an artificial lipid bilayer. However, F509A formed unstable pores and did not result in significant membrane disruption. This was in contrast to Cry41Aa which produced stable pores which resulted in permanent damage to cell membrane and eventual cell lysis. The findings conclude that Cry41Aa is likely to have a 3-domain structure similar to that

of insecticidal is 3-domain Cry toxins. Residue substitutions have indicated that domain II loops of Cry41Aa, like domain II insecticidal toxins, play a role its specificity and ability to exert toxicity on susceptible cells.

Similar to insecticidal Cry toxin domain II loops of Cry41Aa can be explored to enhance toxicity as shown by the more potent mutant Y514A. Cry41Aa is thought to have a mode of action similar to insecticidal 3-domain toxins where it forms pores in the membrane of susceptible cells causing osmotic shock and the eventual lysis of cell (Krishnan *et al*., 2017) .

7.2 Discussion

The research question in this study aimed to investigate the specificity of the Cry41Aa toxin by exploring its domain II loops. The study began by collating literature on the domain II loops which had been identified from revealed crystal structures (Adang *et al*., 2014; Xu *et al*., 2014). The study employed bioinformatic tools that used information from secondary and tertiary protein structures to propose a three-dimensional model of Cry41Aa. It predicted Cry41Aa to have a typical 3-domain globular structure (Krishnan *et al*., 2017). Furthermore, the exposed loops of domain II were predicted and aligned with loops from the revealed Cry crystal structures (Xu *et al*., 2014).

Previous studies on 3-domain Cry toxins have demonstrated toxin specificity by modifying identified specificity regions such as domain II loops to create novel toxins or increase effectiveness of a toxin. (Pigott *et al*., 2008; Abdullah *et al*., 2003; Abdullah and Dean, 2004; Abdul-Rauf and Ellar, 1999).

The first mutagenesis approach attempted to identify the loops responsible for the specificity of Cry41Aa by exchanging loop 3 of Cry41Aa with loop 3 from an insecticidal 3-domain toxin and vice versa. Previous studies on loop exchange have introduced novel toxicities in known Cry toxins confirming that it is indeed these loop that are involved in the specificity of the toxin in question. A study by Abdulla *et al*, (2003) aimed to introduce *Culex* mosquito toxicity in Cry4Ba (toxic to *Anopheles* and *Aedes*) by exchanging loop regions with Cry4Aa into Cry4Ba. Loop 1 and 2 of Cry4Ba were deleted and replaced with loop 1 and 2 of Cry4Aa. These recombinants did not exhibit toxicity to susceptible insects. However, loop 3 insertion of Cry4Aa into Cry4Ba produced a recombinant that gained toxicity to *Culex* (Abdullah *et al.,* 2003, Howlader *et al.,* 2009*)*.

Further investigation by Abdulla and Dean*,* (2004) led to the creation of a Cry4Aa loop1 exchange with loop1 of Cry19Aa. The mutant sustained toxicity to *Culex* (Cry4Aa) and gained toxicity to *Anopheles* and *Aedes* (susceptible to Cry19A). Collated observations of competition binding assay on this loop exchange mutant concluded that there was no difference in its ability to bind BBMV between less or more active recombinant toxins. The study proposed a model whereby nontoxic recombinants bind to BBMV surface receptors but do not induce toxicity. The model suggests that enhanced toxicity

recombinants are able to bind to unique and minor receptors that elicit toxicity (Abdullah and Dean, 2004).

In a study by Zhou *et al.* (2017) the domain II loops of Cry1Ah (toxic to *Helicoverpa armigera* but not *Bombyx mori* larvae) were exchanged with domain II loops of Cry1Ai which is active against Bombyx mori larvae but not *H. armigera*. This mutagenesis approach involved the introduction and exchange of domain II loops 2 and 3 between Cry1Ai and Cry1Ah.

Zhou *et al*. (2017) intended to explore the extent that these loops are involved in the toxins' specificity to their target insect cells. They observed that the Cry1Ai-h-loop 2-exchange mutant, containing loop 2 from Cry1Ah, showed significant toxicity towards *H. armigera*, whilst the Cry1Ah-i-loop 2-exchange mutant containing loop 2 from Cry1Ai did not indicate any toxicity towards H. *armigera*. The loop 2 exchange in Cry1Ah caused loss of toxicity to its target insect *H. armigera*. Similarly, the introduction of loop 3 from Cry1Ai into Cry1Ah also induced the loss of toxicity to *H. armigera* but neither loop 2 or 3 exchange allowed Cry1Ah to gain a new toxicity towards *B. mori*. Contrary to this, is that loop 2 exchange in Cry1Ai where the presence of both loops 2 and 3 of Cry1Ah allowed Cry1Ai to gain toxicity to *H. armigera* which was not exhibited earlier (Zhou *et al*., 2017).

The recombinant Cry41Aa loop 3 of Cry1Ac and Cry41Ab loop 3 of Cry1Ac exchange mutants were not stable in trypsin and did not result in a proteolytic protease core. The

insecticidal Cry1Ac loop 3 of Cry41Aa exchange did result in a proteolytic stable core, however, cell assays indicated that it did not gain toxicity to the HepG2 cell line. We did not assess whether this mutant lost its native toxicity to *M. sexta* or other known target lepidopteran insects as a result of replacing of its native loop 3 with that of Cry41Aa. However, the lack of proteolytic stable mutants of Cry41Aa led to the second mutagenesis approach.

Research has shown that domain swapping has led to new specificities and improved toxicities. In addition to the domain II loops, domain II and III have been implicated in specificity and receptor recognition (Lee *et al*., 1999; Bravo *et al*., 2007; de Maagd *et al*., 1996; Dean *et al*., 1996; Gomez *et al*., 2006; Herrero *et al*., 2004).

A study by de Maagd *et al,* (1996) carried out a study on domain I, II, and III of Cry1Ab and Cry1Ca, they created a number of hybrids by swapping domains from both toxins in an attempt to improve the toxicity towards *Spodoptera exigua* larvae. They observed that combinations of different domain I and II were unusually toxic when combined with domain III of Cry1Ca. This observation extended to domain I and II from other 3-domain Cry toxins Cry1Ac, Cry1Ba, Cry1Ea, or Cry1Fa as long as domain III came from Cry1Ca. It highlighted the role that domain III of Cry1Ca played on the toxicity and specificity that it has towards *S. exigua* (de Maagd *et al*., 1996).

In the quest to understand its specificity towards HepG2 cell lines, hybrids of Cry1Ie and Cry42Aa were created with Cry41Aa. There is no record of domain swapping

mutagenesis being carried out in Cry41Aa. A similar study by Masson *et al*. (1994) created an exchange of non-homologous N terminal regions of Cry1Ac with Cry1Ea and determined that the N-terminal region of a toxin is not important in specificity. This study, conducted both *in vivo* and *in vitro* on cultured insect cells and dissociated midgut epithelial cells, concluded that hybrid toxins can exhibit different toxicity and that this is dependent on bioassay systems in place (Masson *et al*., 1994). Despite this, our current hybrid was designed to introduce a region of N terminus (domain I) of Cry41Aa into Cry1Ea. However, both Cry1Ie-Cry41Aa and Cry42Aa-Cry41Aa hybrids did not result in a protease resistant cores upon activation and thus could not be investigated for their toxicity on HepG2 cells.

The third mutagenesis approach involved residue substitution and specifically alanine residue substitutions to probe the domain II exposed loops of Cry41Aa for specificity after failure to obtain protease resistant cores from previous mutagenesis approaches on Cry41Aa. Research has highlighted the role that hydrophobic residue substitutions of domain II loops play on the specificity and toxicity of 3-domain toxins. (Lee *et al*., 1999; Lee *et al*., 2001; Lu *et al*., 1994; Rajamohan *et al*., 1996c; Rajamohan *et al*., 1996b; Roh *et al*., 2009).

In a recent review on the structure of Cry toxins, Xu *et al*. (2014) proposed an alternative classification system of Cry toxins based on toxin mode of action and structural relatedness. This study suggested that Bt toxins be categorised into 3-domain type α pore forming toxins, Cyt toxins type β-pore forming toxins and aerolysin type β pore

forming toxins (Xu *et al*., 2014). Its highlighted findings on the structurally diverse domain II Cry toxins and highlighted the presence of aromatic residues and their rich presence on the surface of domain II.

The surface exposed aromatic rings have been implicated in binding to polysaccharides and are often found in carbohydrate-protein complexes. Parasporin 2 (Cry46Aa1) has a number of aromatic residues on the surface and these may play a role binding. Although the effects of aromatic residues have been explored in insecticidal Cry toxins, parasporin have remained behind in this aspect. The presence of aromatic residues of the surface of toxins suggests that they may have specific interaction with susceptible cell membranes (Xu *et al*., 2014).

Studies suggest that hydrophobic residues may be responsible for interactions between the cell membranes of target insect cells (Wu *et al*., 2000). Three hydrophobic residues were identified in loop 3 of Cry41Aa and were substituted with alanine F509A, W511A, and Y514A. All three resulted in protease resistant peptides that were purified and their toxicity to HepG2 was investigated.

A 24 h CellTiter-Blue assay was applied to measure the percentage of HepG2 cell viability relative to the buffer. CellTox Green Cell Assay was used to assess HepG2 membrane damage caused by Cry41Aa and the nontoxic F509A. It was concluded that Cry41Aa caused pores in the membrane of HepG2 cells where the small green fluorescent molecule was able to bind to nuclear material. There was no record of any nuclear

binding of the dye molecule in cells incubated with F509A. These findings suggest that F509A did not cause any cell membrane damage even after 28 h of incubation with HepG2 cells.

Cell assay data suggested that the Y514A mutant retained its toxicity to HepG2 cells. W511A was also not toxic to HepG2. These findings highlighted the need to clarify whether it was the type of residue or the position of the residue that lead to the loss of toxicity. This led to further mutagenesis at positions F509 and W511.

A number of degenerate residues were introduced at position F509.and W511 The recombinants F509L, F509S, F50W, and F509Y were made, activated and purified. An attempt to introduce W511Y and W511F residue substitutions were also made. Only the W511F mutant was proteolytically stable and initial cell assays indicated, that together with Y514A, they both retained the cytotoxic properties towards HepG2 cells. Apart from wildtype Cry41Aa, and mutants W511F and Y514A all other mutant did not exhibit cytotoxicity towards HepG2 cell lines.

The loss of toxicity by F509A and W511A could be due to the loss of interaction with HepG2 cells or that interaction was insufficient or ineffective to cause cell death. Initially, western blots were carried out to assess phosphorylation of p38 and activation of the p38 MAP kinase pathway in Cry41Aa and nontoxic F509A. Brasseur *et al.* (2015) investigated the activation of apoptosis by Cry46Aa (Parasporin 2) and observed that Cry46Aa triggered this pathway by phosphorylation of p38 (Brasseur *et al.*, 2015).

Western blots of Cry41Aa and mutant F509A revealed that both were found to induce p38 phosphorylation. These finding suggest that the toxicity of Cry41Aa may be affected by the type of protease used to activate the toxin (Souissi, 2018). It is possible that pore formation by F509A is deficient or insufficient in number to cause cell death. Thus, the type of pore and its effectiveness was investigated by electrophysiology.

Seven cell lines including HL60 were assayed for their viability after incubation with recombinant toxins. Cell assay data concluded that none of the mutants or wildtype Cry41Aa were toxic to cell lines. The HL60 cell line was previously reported by Yamashita *et al.* (2005) to be susceptible to Cry41Aa. In contrast, viability cell assay in this study did not support such a finding, however some studies has shown HL60 cell lines to be susceptible to Cry41Aa when activated by proteinase K as previously used by Yamashita (Souissi, 2018, Domanska, 2016).

A dose response curve for the protein concentration optimised toxins was calculated and the EC_{50} values were determined by probit analysis using SPSS software. Cry41Aa had an EC_{50} value of 4.6 μg/mL, and the mutant Y514A had an EC_{50} value of 2.6 μg/mL. Y514A appeared to be a more potent toxin than Cry41Aa., killing HepG2 cells faster than wildtype Cry41Aa as indicated by microscope observations.

A similar study by Wu *et al*. (2000) on residue substitutions in loop 1 of Cry31Aa also indicated mutants that showed higher toxicity in bioassays of three coleopteran Cry3Aa susceptible species. Here substitution of phenylalanine for tyrosine at positions 350 and

351 affected the dissociation of the toxin which gave rise to enhanced toxicity by mutants. This study proposed that hydrophobic side chains of loop1 residues interact with a target receptor. It further speculated that increased toxicity in phenylalanine mutant substitutions with tyrosine was due to the presence of aromatic phenol rings and not the phenolic hydroxyl group of tyrosine. The study suggested that aromatic residues are involved in hydrophobic interactions with receptors, implicating them in the binding affinity of the toxin to the target receptor. Y514A and W511F retained their toxicity towards HepG2. In mutant Y514A, the aromatic tyrosine residue was substituted with the small non-aromatic alanine residue, yet toxicity appeared higher. HepG2 cells were observed to swell and detach within 10 min of incubation with the mutant. It is relevant to point out that the HepG2 receptors for Cry41Aa are as yet unknown.

An attempt to visualise binding was carried using fluorescein isothiocyanate (FITC). Activated Cry41Aa and F509A were incubated with FITC and then washed with a buffer to remove excess FITC. HepG2 cells were briefly incubated with FITC labelled Cry41Aa and F509A and then washed to remove excess or unbound FITC labelled proteins. The cells were observed under a Zeiss Axiovert 200M microscope using FITC and bright-field channels in a wide-field configuration. Unfortunately, the fluorescence signal was very weak despite the obvious cytotoxic effects of Cry41Aa on Hep G2, there was an absence of fluorescence outlining the cell membrane. Attempts to visualise and measure binding include biotin labelling of Cry41Aa and tagging Cry41Aa with HA-tag. Both of these were unsuccessful in determining the binding of Cry41Aa to HepG2 cells (Domanska, 2016).

Electrophysiology was carried out to investigate if Cry41Aa and the nontoxic F509A form any pores. Whole patch clamping carried out on HepG2 membranes indicated that Cry41Aa does cause pore formation. The pores open in sufficient numbers and are stable enough to cause toxicity and eventual cell death. It also concluded that if pores are formed as a result of HepG2 incubation with F509A, they are not stable enough or sufficient in number. It is possible that the cells are able to recover and a significant change in cell viability or membrane damage would not be observed. PLB single channel activity was applied to study the activity of Cry41Aa and F509A toxin on an artificial membrane. It concluded that Cry41Aa forms stable pores that seem to be voltage dependent and are slightly cationic selective. Although F509A revealed similar findings, the standard error was much higher which indicated the pore type formed was unstable and had different kinetics and opening abilities. Electrophysiology tests indicated that F509A causes pores which were not always stable or effective, hence their inability to induce toxicity.

In this study, a more potent Y514A does not necessarily suggest higher binding affinity to HepG2. The mutant Y514A may induce more pores that are very stable and effective which may explain its higher toxicity when compared to the toxicity of wildtype Cry41Aa. However, without supportive electrophysiology evidence this stands as a speculation. These finding are supported by studies that have implicated aromatic residues in the domain II loop region in initial binding and specificity (Rajamohan *et al*., 1996c; Pacheco *et al*., 2009; Abdullah *et al*., 2003; Dean *et al*., 1996; Lu *et al*., 1994).

Howlader *et al.* (2010) probed the effect of alanine substitution in loop 3 of insecticidal Cry4Aa and observed a lower toxicity to fourth instar *C. pipien mosquitoes* in two mutants: T512A and Y513A. Cry41Aa loop 3 recombinant Y514A was more toxic than the wildtype. This may suggest that a particular residue position is not important to Cry toxins but rather the type of residue at that position. The W511F mutant was also toxic to HepG2 cells. It can be argued that the mutant has retained its toxicity due to the presence the aromatic ring of phenylalanine.

A similar study by Wu *et al.* (2000) suggested that an increase in hydrophobic phenylalanine rings and or the removal of the phenolic hydroxyl groups at some residues in loop binding regions can result in higher binding affinity and subsequently a higher toxicity (Wu *et al.*, 2000).

In addition to the loop 3 mutants, an extra loop alanine substitution was also created as well as degenerate residue substitutions in loop1 of Cry41Aa. The protein concentration for all recombinant loop mutants was optimised with the exception of the following loop1 mutants and loop 3 W511F. Loop1 mutants retained cytotoxicity to HepG2 cells on initial cell assays. These degenerate residue mutants suggested that loop1 may not have a key role in the specificity of Cry41Aa was not probed further.

The W511F was made after the protein optimisation of the loop 3 mutants and was not run on the same SDS PAGE gel at the same time as all investigated mutants. Despite this, measures were taken to apply the Bradford protein assay to measure its protein

concentration before cell assays with HepG2 cell lines. Observations revealed that W511F is toxic to HepG2 cell lines, but its toxicity level compared to wildtype Cry41Aa and the potent Y514A remain unknown.

The extra loop is unique to Cry41Aa, previous attempts to obtain a protease resistant core of Cry41Aa, where the extra loop was completely deleted or partially deleted, were unsuccessful (Banani, 2013; Krishnan, 2013; Krishnan *et al*., 2017). These findings have highlighted the key role that the extra loop may play in the structural integrity of the toxin. All mutagenesis attempts so far have resulted in unstable recombinant proteins that degrade in the solubilisation or activation stage.

An extra loop substitute was created, as a series of alanine substitutions were introduced to create an alanine cassette in the loop. It changed the native secondary structure of the extra loop from sheets to helices. The mutant was proteolytically stable and its concentration was optimised along with the other loop substitution mutants. Cell assays suggest that the extra loop mutant was not toxic to HepG2 cell lines. A clear and definite role for this distinctive loop is yet to be found, although its presence is critical for obtaining a protease resistant mutant.

Studies on the type of protease used to activate Cry toxins can affect their specificity and toxicity. Studies on insecticidal Cry toxins revealed that specificity of a toxin is linked to the type of protease present in the insect gut. Haider *et al*. (1986) proposed that different proteases produce different protease resistant proteins depending of the

amino acid sequence of the toxin(s), and it is the different molecular weights of the activated toxins that can define their specificity. For example, Mizuki *et al*. (2000) demonstrated that Cry31Aa (parasporin 1) had a different specificity when activated by different proteases.

The study demonstrated that Cry31Aa was toxic to cancer cell line MOLT-4 only when activated by Proteinase K or trypsin. Activation by chymotrypsin resulted in proteolysis profiles which differ to those produced by Proteinase K and trypsin and the loss of toxicity towards MOLT-4 cell lines. (Mizuki *et al*, 2000).

Other studies on the parasporins linking their cytotoxicity to their activation have emerged. Brasseur *et al.* (2015) demonstrated that Cry46Aa (parasporin 2) was cytocidal HepG2, MCF-7, KLE, Hec-1A, MDA-MB231 cell lines as well as PC-3 cells when activated with Proteinase K, but no cytotoxicity was observed when activated with trypsin. Yamashita *et al.* (2005) demonstrated that the toxic effect of Cry41Aa towards HepG2 and HL-60 cell lines was dependent on activation by Proteinase K (Yamashita *et al,* 2005). Similar observations were made by Souissi (2018). Findings concluded that Cry41Aa, potentially a pore forming toxin, has a similar mode of action to insecticidal 3-domain toxins (Krishnan *et al.*, 2017).

Recently, some studies emerged that question the safety and risk assessment of Cry toxin. The confirmed discovery of parasporins and their ability to interact with vertebrate cell lines is drawing attention to the risk assessment of Cry toxin. This

previously unreported ability to interact with mammalian cells has highlighted the need for better understanding of Cry toxin specificity and mode of action. Indeed Cry toxins were deemed safe due to their lack of persistence in the environment and narrow toxicity range (de Maagd *et al.*, 2003).

However, a report of transgenic plants expressing Cry toxin has concluded that non-target herbivores and insect predators throughout the food chain can obtain plant *Bt* Cry toxin. These were traced as far as the third trophic level in some cases (Torres *et al.*, 2006). The study recognises that their findings are dependent on the amount of *Bt* transgenic plant eaten by a prey, and how much of this prey is available to and consumed by the predators. Torres *et al.* (2006) concluded that the persistence of plant *Bt* toxins in the food chain, is dependent on community structure and dynamics of crop pest insects and their predators, as well as the availability of alternative *Bt* free prey. Nevertheless, the study highlights the need for more field research and less reliance on laboratory research based on artificial conditions.

Despite the number of studies dedicated to understanding Cry toxin mode of action and the driving factor for their evolutionary diversity, there is not t a consensus on either (Then, 2010). Cry toxins with preferential toxicity to cancer are already being investigated for their potential application as an alternative to the current cancer drugs.

In a study by Wong, *et al.* (2010) which studied the binding affinity of *Bacillus thuringiensis* 18 Toxin to CEM-SS cancer cell (leukaemia) line and whether it competes with commercially available anticancer drugs. The study concluded that there was no completion between Toxin 18 and other drugs tested, furthermore it suggested that Toxin 18 was likely to bind to a different receptor on the surface of CEM-SS cell and thus is thought to have a different mode of action compared to the other drugs used in this study (Wong *et al*., 2010).

There is a long history of the safe use of *Bt* and transgenic *Bt* crops to control agricultural and disease carrying pests. some studies have claimed that *Bt* can increase allergenic potential, or cause hematotoxic reactions, however upon closer investigation the claims did not stand up to scientific scrutiny (Koch *et al.*, 2015).

The supposed toxicity exhibited by *Bt* toxins in particular, to mammalian cells, is debatable and many have argued that the term 'toxic' is an unfitting definition for the effects observed when *Bt* interacts with mammalian cells. However, the interaction, and effects of interactions, are observed and thus require further investigation (Rubio-Infante and Moreno-Fierros, 2016).

There are still some avenues to explore from the Cry41Aa loop substitution mutants. Attempts to use fluorescent microscopy to visualise Cry41Aa or its mutants' interactions with susceptible cancer cell lines proved futile. It is possible that only a small amount of Cry41Aa is needed to cause toxicity of that the receptor that it interacts with is present

in low concentrations in the cell membrane. It's worth exploring the nature of the toxicity of mutants Y514A and W511F in a quest to better understand the specificity of Cry41Aa and to investigate its binding affinity compared to its mutants. A series of alanine substitutions, where substitution occurs in more than one position at the same time, was not created. For example, it is not known whether a mutant with alanine substitution in positions 509, 511, and 514 at the same time would still retain toxicity despite the presence of an alanine at position Y514. Furthermore, loop1 substitutes of only alanine would help to clarify the role that aromatic rings may play in the toxicity of Cry41Aa and whether it has hydrophobic interaction with HepG2 cell membranes.

Future work should include activation of mutants by different proteases such as Proteinase K, and analysis whether these mutants demonstrate a different toxicity after activation with a different protease. Further work on electrophysiology analysis of Cry41Aa treated HL60 cell lines would reveal the characteristics of pores, if any are formed as was the case with F509A mutant. single channel electrophysiology can indicate the stability of pores, frequency, opening probability and duration of opening of pores formed by toxic mutant W511F, and Y514A. Future work should attempt to identify receptor(s) of Cry41Aa in susceptible cell lines in order to gain a better understanding its mode of action and how it can potentially contribute to the treatment of cancer.

8.0 Bibliography

Abdul-Rauf, M. and Ellar, D. J. (1999) 'Mutations of loop 2 and loop 3 residues in domain II of *Bacillus thuringiensis* Cry1C delta-endotoxin affect insecticidal specificity and initial binding to *Spodoptera littoralis* and *Aedes aegypti* midgut membranes', *Current Microbiology,* 39(2), pp. 94-98.

Abdullah, M. A. F., Alzate, O., Mohammad, M., McNall, R. J., Adang, M. J. and Dean, D. H. (2003) 'Introduction of *Culex* toxicity into *Bacillus thuringiensis* Cry4Ba by protein engineering', *Applied and Environmental Microbiology,* 69(9), pp. 5343-5353.

Abdullah, M. A. F. and Dean, D. H. (2004) 'Enhancement of Cry19Aa mosquitocidal activity against *Aedes aegypti* by mutations in the putative loop regions of domain II', *Applied and Environmental Microbiology,* 70(6), pp. 3769-3771.

Adalat, R., Saleem, F., Crickmore, N., Naz, S. and Shakoori, A. (2017) *In Vivo* Crystallization of three-domain Cry toxins', *Toxins*, 9(3), pp 80. doi:10.3390/toxins9030080.

Adang, M. J., Crickmore, N. and Jurat-Fuentes, J. L. (2014) 'Diversity of *Bacillus thuringiensis* crystal toxins and mechanism of action', *Advances in insect physiology*: *Elsevier*, pp. 39-87.

Agaisse, H. and Lereclus, D. (1995) 'How does *Bacillus thuringiensis* produce so much insecticidal crystal protein?', *Journal of bacteriology,* 177(21), pp. 6027-6032.

Akiba, T., Abe, Y., Kitada, S., Kusaka, Y., Ito, A., Ichimatsu, T., Katayama, H., Akao, T., Higuchi, K., Mizuki, E., Ohba, M., Kanai, R. and Harata, K. (2009) 'Crystal structure of the parasporin-2 *Bacillus thuringiensis* toxin that recognizes cancer cells', *Journal of molecular biology,* 386(1), pp. 121-133.

Akiba, T., Ichimatsu, T., Katayama, H., Akao, T., Nakamura, O., Mizuki, E., Ohba, M. and Harata, K. (2005) 'Structure of parasporin-1, a novel bacterial cytotoxin against human cancer cells', *Journal of biochemistry*, 137(1), pp. 17-25.

Akiba, T. and Okumura, S. (2016) 'Parasporins 1 and 2: their structure and activity', *Journal of invertebrate pathology*, 142, pp. 44-49. doi: 10.1016/j.jip.2016.10.005.

Angus, T. A. (1954) 'A bacterial toxin paralysing silkworm larvae', *Nature,* 173(4403), pp. 545-6.

Arenas, I., Bravo, A., Soberón, M. and Gómez, I. (2010) 'Role of alkaline phosphatase from *Manduca sexta* in the mechanism of action of *Bacillus thuringiensis* Cry1Ab Toxin', *Journal of Biochemistry,* 285(17), pp. 12497-503.

Atsumi, S., Inoue, Y., Ishizaka, T., Mizuno, E., Yoshizawa, Y., Kitami, M. and Sato, R. (2008) 'Location of the *Bombyx mori* 175 KDa cadherin-like protein-binding site on *Bacillus thuringiensis* Cry1Aa toxin', *Febs Journal,* 275(19), pp. 4913-4926.

Banani, F. 2013. Biological Function of the extra loop of Cry41Aa toxin. University of Sussex.

Baum and Malvar (2005) 'Regulation of insecticidal crystal protein production in *Bacillus thuringiensis*', *Molecular Microbiology,* 18, pp. 1-12.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L. and Schwede, T. (2014) 'SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information', *Nucleic Acids Research*, 42 (Web Server issue), pp. 252-258.

Bietlot, H. P., Vishnubhatla, I., Carey, P. R., Pozsgay, M. and Kaplan, H. (1990) 'Characterization of the cysteine residues and disulphide linkages in the protein crystal of *Bacillus thuringiensis*', *The Biochemical journal,* 267(2), pp. 309-315.

Boonserm, P., Davis, P., Ellar, D. J. and Li, J. (2005) 'Crystal structure of the mosquito-larvicidal toxin Cry4Ba and Its biological implications', *Journal of Molecular Biology,* 348(2), pp. 363-382.

Boonserm, P., Mo, M., Angsuthanasombat, C. and Lescar, J. (2006) 'Structure of the functional form of the mosquito larvicidal Cry4Aa toxin from *Bacillus thuringiensis* at a 2.8-angstrom resolution', *Journal of Bacteriology,* 188(9), pp. 3391-3401.

Brasseur, K., Auger, P., Asselin, E., Parent, S., Cote, J. C. and Sirois, M. (2015) 'Parasporin-2 from a new *Bacillus thuringiensis* 4R2 strain induces caspases activation and apoptosis in human cancer cells', *PLoS One,* 10(8), pp. 106-135.

Bravo, A. (1997) 'Phylogenetic relationships of *Bacillus thuringiensis* delta-endotoxin family proteins and their functional domains', *Journal of Bacteriology,* 179(9), pp. 2793-2801.

Bravo, A., Gill, S. S. and Soberón, M. (2007) 'Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control', *Toxicon,* 49(4), pp. 423-35.

Bravo, A., Gomez, I., Conde, J., Munoz-Garay, C., Sanchez, J., Miranda, R., Zhuang, M., Gill, S. S. and Soberon, M. (2004) 'Oligomerization triggers binding of a *Bacillus thuringiensis* Cry1Ab pore-forming toxin to aminopeptidase N receptor leading to insertion into membrane microdomains', *Biochim Biophys Acta,* 1667(1), pp. 38-46.

Bravo, A., Gómez, I., Porta, H., García-Gómez, B. I., Rodriguez-Almazan, C., Pardo, L. and Soberón, M. (2013) 'Evolution of δ *Bacillus thuringiensis* Cry toxins insecticidal activity: evolution of *Bt* toxins', *Microbial Biotechnology,* 6(1), pp. 17-26.

Bravo, A., Likitvivatanavong, S., Gill, S. S. and Soberón, M. (2011) 'Bacillus thuringiensis: a story of a successful bioinsecticide', *Special Issue: Toxicology and resistance,* 41(7), pp. 423-431.

Bravo, A., S. Gill, S. and Soberón, M. (2018) '*Bacillus thuringiensis : Mechanisms and use' Comprehensive molecular insect science,* Elsevier, pp. 247-278.

Bravo, A. and Soberon, M. (2008) 'How to cope with insect resistance to *Bt* toxins?', *Trends in biotechnology,* 26(10), pp. 573-579.

Brasseur, K., Auger, P., Asselin, E., Parent, S., Cote, J.-C. and Sirois, M. (2015). Parasporin-2 from a new *Bacillus thuringiensis* 4R2 strain induces caspases activation and apoptosis in human cancer cells. *Plos One*, 10.

Cancino-Rodezno, A., Alexander, C., Villaseñor, R., Pacheco, S., Porta, H., Pauchet, Y., Soberón, M., Gill, S. S. and Bravo, A. (2010) 'The mitogen-activated protein kinase p38 is involved in insect defense against Cry toxins from *Bacillus thuringiensis*', *Insect Biochemistry and Molecular Biology,* 40(1), pp. 58-63.

Caramori, T., Albertini, A. M. and Galizzi, A. (1991) 'In vivo generation of hybrids between two *Bacillus thuringiensis* insect-toxin-encoding genes', *Gene,* 98(1), pp. 37-44.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003) 'Multiple sequence alignment with the Clustal series of programs', Nucleic acids research, 31(13), pp. 3497-3500.

Chikawa, Uemori, Yasutake, Kagoshima, Mizuki and Ohba (2008) 'Failure to phenotypically discriminate between non-insecticidal *Bacillus thuringiensis* strains with anticancer parasporins (PS2, PS3, and PS4) *Bacillus thuringiensis* strains that produce insecticidal Cry proteins', *Applied Entomolology ang Zoolology,* 43, pp. 421-426

Chougule, N. P., Li, H., Liu, S., Linz, L. B., Narva, K. E., Meade, T. and Bonning, B. C. (2013) 'Retargeting of the δ *Bacillus thuringiensis* toxin Cyt2Aa against hemipteran insect pests', *Proceedings of the National Academy of Sciences,* 110(21), pp. 8465.

Crickmore, Zeigler, Schnepf, Rie, V., Lereclus, Bravo, B. and Dean (2018) *Bacillus thuringiensis toxin nomenclature*. Available at: http://www.lifesci.sussex.ac.uk/Home/Neil_Crickmore/Bt/.

Crickmore, N., Bone, E. J., Williams, J. A. and Ellar, D. J. (1995) 'Contribution of the individual components of the δ-endotoxin crystal to the mosquitocidal activity of *Bacillus thuringiensis* subsp. *israelensis*', *FEMS Microbiology Letters,* 131(3), pp. 249-254.

Crickmore, N. and Ellar, D. J. (1992) 'Involvement of a possible chaperonin in the efficient expression of a cloned CryllA δ endotoxin gene in *Bacillus thuringiensis*', *Molecular Microbioiogy* 6, pp. 1533-1537.

Crickmore, N., Wheeler, V. and Ellar, D. (1994) 'Use of an operon fusion to induce expression and crystallisation of a *Bacillus thuringiensis* endotoxin encoded by a cryptic gene', *Molecular General Genetics,* 242, pp. 365-368.

Crickmore, N., Zeigler, D. R., Feitelson, J., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J. and Dean, D. H. (1998) 'Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins', *Microbiology and Molecular Biology Reviews,* 62(3), pp. 807.

De Maagd, R., Bravo, A. and Crickmore, N. (2005) *Bt toxin not guilty by association.nature biotechnology,* 23(7), pp. 791.

de Maagd, R. A., Bravo, A. and Crickmore, N. (2001) 'How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world', *Trends in Genetics,* 17(4), pp. 193-199.

de Maagd, R. A., Kwa, M. S., van der Klei, H., Yamamoto, T., Schipper, B., Vlak, J. M., Stiekema, W. J. and Bosch, D. (1996) 'Domain III substitution in *Bacillus thuringiensis* delta-endotoxin CryIA(b) results in superior toxicity for

Spodoptera exigua and altered membrane protein recognition', *Applied and Environmental Microbiology,* 62(5), pp. 1537.

de Maagd, R. A., Weemen-Hendriks, M., Stiekema, W. and Bosch, D. (2000) '*Bacillus thuringiensis* delta-endotoxin Cry1C domain III can function as a specificity determinant for *Spodoptera exigua* in different, but not all, Cry1-Cry1C hybrids', *Applied and Environmental Microbiology,* 66(4), pp. 1559-1563.

Dean, D. H., Rajamohan, F., Lee, M. K., Wu, S. J., Chen, X. J., Alcantara, E. and Hussain, S. R. (1996) 'Probing the mechanism of action of *Bacillus thuringiensis* insecticidal proteins by site-directed mutagenesis - A minireview', *Gene,* 179(1), pp. 111-117.

Dean. H, D. and Sylvis Liu, X. (2006) 'Redesigning *Bacillus thuringiensis* Cry1Aa toxin into a mosquito toxin', *Protein Engineering Design and Selection,* 19, pp. 107-111.

deMaagd, Bravo, A., Berry, C., Crickmore, N. and Schnepf, E. (2003a) 'Structure, diversity, and evolution of protein toxins from spore froming entomopathogenic bacteria ', *Annual Review of Genetics,* 37, pp. 409-433.

Domanska, B. (2016) *Mode of action of a human cancer cell active toxin (parasporin-3) from Bacillus thuringiensis.* University of Sussex.

Du, C., Martin, P. A. W. and Nickerson, K. W. (1994) 'Comparison of disulfide contents and solubility at alkaline ph of insecticidal and noninsecticidal *Bacillus thuringiensis* protein crystals', *Applied environmental microbiology,* 60(10), pp. 3847-53.

Du, C. and Nickerson, K. W. (1996) '*Bacillus thuringiensis* Hd-73 spores have surface-localized Cry1Ac toxin: physiological and pathogenic consequences', *Applied and environmental microbiology,* 62(10), pp. 3722-3726.

Ekino, K., Okumura, S., Ishikawa, T., Kitada, S., Saitoh, H., Akao, T., Oka, T., Nomura, Y., Ohba, M., Shin, T. and Mizuki, E. (2014) 'Cloning and characterization of a unique cytotoxic protein parasporin-5 produced by *Bacillus thuringiensis* A1100 strain', *Toxins (Basel),* 6(6), pp. 1882-95.

Florez, A. M., Osorio, C. and Alzate, O. (2012) 'Protein engineering of *Bacillus thuringiensis* δ-endotoxins', in Sansinenea, E. (ed.) *Bacillus thuringiensis Biotechnology*. Dordrecht: Springer Netherlands, pp. 93-113.

Fujii, Y., Tanaka, S., Otsuki, M., Hoshino, Y., Morimoto, C., Kotani, T., Harashima, Y., Endo, H., Yoshizawa, Y. and Sato, R. (2013) 'Cry1Aa binding to the cadherin receptor does not require conserved amino acid sequences in the domain II loops', *Bioscience Reports,* 33, pp. 103-112.

Gabriel Narvaez, Dong Xu, Jean-Charles C Ô te and Schwartz, J.-L. 'Parasporin PS1Aa2 induces ionic channels in lipid bilayer membranes and calcium oscillations in sensitive cells', *47 [th] Annual Meeting of the Society for Invertebrate Pathology.*, Maniz, Germany, 2014/09/11, 60-60.

Gahan, L. J., Gould, F. and Heckel, D. G. (2001) 'Identification of a gene associated with *Bt* resistance in *Heliothis virescens*', *Science,* 293(5531), pp. 857.

Galitsky, N., Cody, V., Wojtczak, A., Ghosh, D., Luft, J. R., Pangborn, W. and English, L. (2001) 'Structure of the insecticidal bacterial delta-endotoxin Cry3Bb1 of

Bacillus thuringiensis', *Acta* Cry*stallographica. Section D, Biological Cry*stallography*, 57(Pt 8), pp. 1101-1109.

García-Robles, I., Ochoa-Campuzano, C., Sánchez, J., Contreras, E., Real, M. D. and Rausell, C. (2012) 'Functional significance of membrane associated proteolysis in the toxicity of Bacillus thuringiensis Cry3Aa toxin against Colorado potato beetle', *Toxicon*, 60(6), pp. 1063-1071.

George, Z. O. (2011). Expression and genetic manipulation of Bacillus thuringiensis toxins for improved toxicity and development of a protocol for in vivo selection of toxin variants with improved activity. PhD, University of Sussex.

Glick, B. R. and Pasternak, J. J. (1998) '*Molecular Biotechology'.* 22nd edition*. Washington, D.C, *American society for microbiology press*, pp. 257-270.

Goldman, D. E. (1943) 'Potential, impedance, and rectification in membranes'. *Journal of general physiology*, 27(1),pp 37-60.

Gomez, I., Arenas, I., Benitez, I., Miranda-Rios, J., Becerril, B., Grande, R., Almagro, J. C., Bravo, A. and Soberon, M. (2006) 'Specific epitopes of domains II and III of *Bacillus thuringiensis* Cry1Ab toxin involved in the sequential interaction with cadherin and aminopeptidase-N receptors in *Manduca sexta*', *J Biol Chem,* 281(45), pp. 34032-9.

Gomez, I., Dean, D. H., Bravo, A. and Soberon, M. (2003) 'Molecular basis for *Bacillus thuringiensis* Cry1Ab toxin specificity: two structural determinants in the *Manduca sexta* Bt-R1 receptor interact with loops alpha-8 and 2 in domain II of Cy1Ab toxin', *Biochemistry,* 42(35), pp. 10482-9.

Gomez, I., Miranda-Rios, J., Rudino-Pinera, E., Oltean, D. I., Gill, S. S., Bravo, A. and Soberon, M. (2002) 'Hydropathic complementarity determines interaction of epitope (869)HITDTNNK(876) in *Manduca sexta* Bt-R(1) receptor with loop 2 of domain II of *Bacillus thuringiensis* Cry1A toxins', *Journal of Biological Chemistry,* 277(33), pp. 30137-43.

Gomez, I., Pardo-Lopez, L., Munoz-Garay, C., Fernandez, L. E., Perez, C., Sanchez, J., Soberon, M. and Bravo, A. (2007) 'Role of receptor interaction in the mode of action of insecticidal Cry and Cyt toxins produced by *Bacillus thuringiensis*', *Peptides,* 28(1), pp. 169-173.

Gonzalez, E., Granados, J. C., Short, J. D., Ammons, D. R. and Rampersad, J. (2011) 'Parasporins from a Caribbean Island: evidence for a globally dispersed *Bacillus thuringiensis* strain', *Current Microbiology,* 62(5), pp. 1643-8.

Grochulski, P., Masson, L., Borisova, S., Pusztai-Carey, M., Schwartz, J. L., Brousseau, R. and Cygler, M. (1995) '*Bacillus thuringiensis* CryIA(a) insecticidal toxin: crystal structure and channel formation', *Journal of Molecular Biology,* 254(3), pp. 447-464.

Guo, S., Ye, S., Liu, Y., Wei, L., Xue, J., Wu, H., Song, F., Zhang, J., Wu, X., Huang, D. and Rao, Z. (2009) 'Crystal structure of *Bacillus thuringiensis* Cry8Ea1: An insecticidal toxin toxic to underground pests, the larvae of *Holotrichia parallela*', *Journal of Structural Biology,* 168(2), pp. 259-266.

Güereca, L. and Bravo, A. (1999) 'The oligomeric state of *Bacillus thuringiensis* Cry toxins in solution', *Biochim biophys acta,* 1429(2), pp. 342-50.

Haider, M. Z., Knowles, B. H. & Ellar, D. J. (1986). Specificity of *Bacillus-thuringiensis* var colmeri insecticidal delta-endotoxin is determined by differential proteolytic processing of the protoxin toxin by larval gut proteases, *European journal of biochemistry*, 156, pp.531-540.

Herrero, S., González-Cabrera, J., Ferré, J., Bakker, P. L. and de Maagd, R. A. (2004) 'Mutations in the *Bacillus thuringiensis* Cry1Ca toxin demonstrate the role of domains II and III in specificity towards *Spodoptera exigua* larvae', *The Biochemical iournal,* 384(Pt 3), pp. 507-513.

Höfte, H. and Whiteley, H. R. (1989) 'Insecticidal crystal proteins of *Bacillus thuringiensis'*, *Microbiological reviews,* 53(2), pp. 242-255.

Howlader, M. T. H., Kagawa, Y., Miyakawa, A., Yamamoto, A., Taniguchi, T., Hayakawa, T. and Sakai, H. (2010) 'Alanine scanning analyses of the three major loops in domain II of *Bacillus thuringiensis* mosquitocidal toxin Cry4Aa', *Applied and environmental microbiology,* 76(3), pp. 860-865.

Howlader, M. T. H., Kagawa, Y., Sakai, H. and Hayakawa, T. (2009) 'Biological properties of loop-replaced mutants of *Bacillus thuringiensis* mosquitocidal Cry4Aa', *Journal of bioscience and bioengineering,* 108(3), pp. 179-183.

Huffman, D. L., Abrami, L., Sasik, R., Corbeil, J., van der Goot, F. G. and Aroian, R. V. (2004) 'Mitogen-activated protein kinase pathways defend against bacterial pore-forming toxins', *Proceedings of the national academy of sciences of the united states of America,* 101(30), pp. 10995.

Hui, F., Scheib, U., Hu, Y., Sommer, R. J., Aroian, R. V. and Ghosh, P. (2012) 'Structure and glycolipid binding properties of the nematicidal protein Cry5B', *Biochemistry,* 51(49), pp. 9911-9921.

Ito Akio, Sasaguri Yasuyuki, Kitada Sakae, Kusaka Yoshitomo, Kuwano Kyoko, Masutomi Kenjiro, Mizuki Eiichi, Akao Tetsuyuki and Michio, O. (2004) 'A *Bacillus thuringiensis* crystal protein with selective cytocidal action to human cells', *The Journal of Biological chemistry,* 14, pp. 21282-21286.

Jez, J. M. (2017) 'Revisiting protein structure, function, and evolution in the genomic era', *Journal of invertebrate pathology,* 142, pp. 11-15.

Jiménez-Juárez, N., Muñoz-Garay, C., Gómez, I., Gill, S. S., Soberón, M. and Bravo, A. (2008) 'The pre-pore from *Bacillus thuringiensis* Cry1Ab toxin is necessary to induce insect death in *Manduca sexta*', *Peptides,* 29(2), pp. 318-23.

Jurat-Fuentes, J. L. and Adang, M. J. (2004) 'Characterization of a Cry1Ac-receptor alkaline phosphatase in susceptible and resistant *Heliothis virescens* larvae', *European journal of biochemistry,* 271(15), pp. 3127-3135.

Jurat-Fuentes, J. L. and Adang, M. J. (2006) 'Cry toxin mode of action in susceptible and resistant *Heliothis virescens* larvae', *Journal of invertebrate pathology,* 92(3), pp. 166-171.

Jurat-Fuentes, J. L. and Adang, M. J. (2007) 'A proteomic approach to study Cry1Ac binding proteins and their alterations in resistant *Heliothis virescens* larvae', *Journal of iInvertebrate pathology,* 95(3), pp. 187-91.

Jurat-Fuentes, J. L. and Crickmore, N. (2017) 'Specificity determinants for Cry insecticidal proteins: Insights from their mode of action', *Journal of invertebrate pathology,* 142, pp. 5-10.

Katayama, Kusaka, Yokota, Akao, Kojima, Nakamura, Mekada and Mizuki (2005) 'Parasporin-1, a novel cytotoxic protein from *Bacillus thuringiensis*, Induces Ca2 Influx and a sustained elevation of the cytoplasmic Ca2 concentration in toxin-sensitive Cells', *The Journal of biological chemistry*, 282 (10), pp. 7742-7752.

Katayama, H., Kusaka, Y., Mizuki, E., Government, F. P. and Prefecture, F. (2011) *Parasporin-1 receptor and use thereof*. Google. [Online]. Available at: https://patents.google.com/patent/EP2273266A1/en.

Kitada, Abe, Ito, Kuge, Akao, Mizuki and Ohba 'Molecular identfication and cytocidal action of parasporin, a protein group of novel crystal toxins targeting human cancer cells', *6th Pacific conference on the biotechnology of Bacillus thuringiensis and its environmental impact*, 2005, Victoria, 23-30.

Kitada, S., Abe, Y., Shimada, H., Kusaka, Y., Matsuo, Y., Katayama, H., Okumura, S., Akao, T., Mizuki, E., Kuge, O., Sasaguri, Y., Ohba, M. and Ito, A. (2006) 'Cytocidal actions of parasporin-2, an anti-tumor crystal toxin from Bacillus thuringiensis', *The Journal of biological chemistry,* 281(36), pp. 26350-26360.

Knowles, B. H. and Dow, J. A. T. (1993) 'The crystal δ-endotoxins of *Bacillus thuringiensis*: models for their mechanism of action on the insect gut', *BioEssays,* 15(7), pp. 469-476.

Knowles, B. H. and Ellar, D. J. (1987) 'Colloid-osmotic lysis is a general feature of the mechanism of action of *Bacillus thuringiensis* δ-endotoxins with different insect specificity', *Biochimica et biophysica acta (BBA) - general subjects,* 924(3), pp. 509-518.

Koch, M. S., Ward, J. M., Levine, S. L., Baum, J. A., Vicini, J. L. and Hammond, B. G. (2015) 'The food and environmental safety of *Bt* crops', *Frontiers in plant science,* 6, pp. 283-283.

Krishnamoorthy, M., Jurat-Fuentes, J. L., McNall, R. J., Andacht, T. and Adang, M. J. (2007) 'Identification of novel Cry1Ac binding proteins in midgut membranes from *Heliothis virescens* using proteomic analyses', *Insect biochemstry and molecular biology,* 37(3), pp. 189-201.

Krishnan, V. (2013) *Investigation of parasporins, the cytotoxic proteins from the bacterium Bacillus thuringiensis.* Doctoral, University of Sussex.

Krishnan, V., Domanska, B., Elhigazi, A., Afolabi, F., West, M. J. and Crickmore, N. (2017) 'The human cancer cell active toxin Cry41Aa from B*acillus thuringiensis* acts like its insecticidal counterparts', *The biochemical Journal,* 474(10), pp. 1591-1602.

Lee, M. K., Jenkins, J. L., You, T. H., Curtiss, A., Son, J. J., Adang, M. J. and Dean, D. H. (2001) 'Mutations at the arginine residues in alpha8 loop of *Bacillus thuringiensis* delta-endotoxin Cry1Ac affect toxicity and binding to *Manduca sexta* and *Lymantria dispar* aminopeptidase N', *FEBS letters,* 497(2-3), pp. 108-112.

Lee, M. K., You, T. H., Curtiss, A. and Dean, D. H. (1996) 'Involvement of two amino acid residues in the loop region of *Bacillus thuringiensis* Cry1Ab toxin in toxicity and

binding to *Lymantria dispar*', *Biochemical and Biophysical Research Communications,* 229(1), pp. 139-146.

Lee, M. K., You, T. H., Gould, F. L. and Dean, D. H. (1999) 'Identification of residues in domain III of *Bacillus thuringiensis* Cry1Ac toxin that affect binding and toxicity', *Applied and Environmental Microbiology,* 65(10), pp. 4513-4520.

Li, J. D., Carroll, J. and Ellar, D. J. (1991) 'Crystal structure of insecticidal delta-endotoxin from *Bacillus thuringiensis* at 2.5 A resolution', *Nature,* 353(6347), pp. 815-821.

Likitvivatanavong, S., Aimanova, K. G. and Gill, S. S. (2009) 'Loop residues of the receptor binding domain of *Bacillus thuringiensis* Cry11Ba toxin are important for mosquitocidal activity', *FEBS Letters,* 583(12), pp. 2021-2030.

Lilien, J. and Balsamo, J. (2005) 'The regulation of cadherin-mediated adhesion by tyrosine phosphorylation/dephosphorylation of β-catenin', *Current Opinion in Cell Biology,* 17(5), pp. 459-465.

Liu, X., Peng, D., Luo, Y., Ruan, L., Yu, Z. and Sun, M. (2009) 'Construction of an *Escherichia coli* to *Bacillus thuringiensis* shuttle vector for large DNA fragments', *Applied Microbiology Biotechnology,* 82, pp. 765-772.

López-Pazos, S. A., Rojas Arias, A. C., Ospina, S. A. and Cerón, J. (2010) 'Activity of *Bacillus thuringiensis* hybrid protein against a *lepidopteran* and a *coleopteran* pest', *FEMS Microbiology letters,* 302(2), pp. 93-98.

Lu, H., Rajamohan, F. and Dean, D. H. (1994) 'Identification of amino-acid-residues of *Bacillus-thuringiensis* delta-endotoxin Cry1Aa associated with membrane-binding and toxicity to *Bombyx-mori*', *Journal of Bacteriology,* 176(17), pp. 5554-5559.

Lucena, W. A., Pelegrini, P. B., Martins-de-Sa, D., Fonseca, F. C. A., Gomes, J. E., de Macedo, L. L. P., da Silva, M. C. M., Oliveira, R. S. and Grossi-de-Sa, M. F. (2014) 'Molecular approaches to improve the insecticidal activity of *Bacillus thuringiensis* Cry toxins', *Toxins (Basel),* 6(8), pp. 2393-423.

Mahadeva Swamy, H. M., Asokan, R. and Mahmood, R. (2014) 'Insilico structural 3d modelling of novel Cry1Ib9 and Cry3A toxins from local isolates of *Bacillus thuringiensis*', *Indian journal of Microbiology,* 54(1), pp. 94-103.

Masson, L., Mazza, A., Gringorten, L., Baines, D., Aneliunas, V. and Brousseau, R. (1994) 'Specificity domain localization of *Bacillus-thuringiensis* insecticidal toxins is highly dependent on the bioassay system', *Molecular microbiology,* 14(5), pp. 851-860.

Mathew and Verma (2009) 'Humanized immunotoxins: a new generation of immunotoxins for targeted cancer therapy', *Cancer Science*, 100, pp. 1359-1365, Available: Wiley. DOI: 10.1111/j.1349-7006.2009.001192.x (Accessed 20/08/2010).

McGuffin, L. J., Bryson, K. and Jones, D. T. (2000) 'The PSIPRED protein structure prediction server', Bioinformatics, 16(4), pp. 404-405.

McNall, R. J. and Adang, M. J. (2003) 'Identification of novel *Bacillus thuringiensis* Cry1Ac binding proteins in *Manduca sexta* midgut through proteomic analysis', *Insect Biochemistry and molecular biology,* 33(10), pp. 999-1010.

Mesnage, R., Clair, E., Gress, S., Then, C., Szekacs, A. and Seralini, G. E. (2013) 'Cytotoxicity on human cells of Cry1Ab and Cry1Ac Bt insecticidal toxins alone or with a glyphosate-based herbicide', *Journal of applied toxicology,* 33(7), pp. 695-9.

Mizuki, Ohba, Akao, Yamashita, Saitoh and Park (1999) 'Unique activity associated with non-insecticidal *Bacillus thuringiensis* parasporal inclusions: *in vitro* cell-killing action on human cancer cells', *Journal of Applied Microbiology* 86, pp. 477-486.

Mizuki, Park, Saitoh, Yamashita, Aakao, Higughi and Ohba (2000) 'Parasporin, a human leukemic cell-recognizing parasporal protein of *Bacillus thuringiensis*', *American society for microbiology,* 7,(4) pp. 625-634.

Moar, W. J., Evans, A. J., Kessenich, C. R., Baum, J. A., Bowen, D. J., Edrington, T. C., Haas, J. A., Kouadio, J.-L. K., Roberts, J. K., Silvanovich, A., Yin, Y., Weiner, B. E., Glenn, K. C. and Odegaard, M. L. (2017) 'The sequence, structural, and functional diversity within a protein family and implications for specificity and safety: The case for ETX_MTX2 insecticidal proteins', *Journal of Invertebrate Pathology,* 142, pp. 50-59.

Morse, R. J., Yamamoto, T. and Stroud, R. M. (2001) 'Structure of Cry2Aa suggests an unexpected receptor binding epitope', *Structure (London, England: 1993),* 9(5), pp. 409-417.

Nachimuthu, S. and Polumetla Ananda, K. (2004) 'Protein engineering of δ-endotoxins of Bacillus thuringiensisa', *Electronic journal of biotechnology; Vol 7, No 2 (2004).*

Nadarajah, V. D., Chai, S. H., Mohammed, S. M., Chan, K. K. and Kanakeswary, K. (2006) 'Malaysian mosquitocidal soil bacterium (*Bacillus thuringiensis*) strains with selective hemolytic and lectin activity against human and rat erythrocytes', *Southeast Asian Journal of tropical medicine and public health,* 37(1), pp. 67-78.

Nagamatsu, Okamura, Saitoh, Akao and Mizuki (2010) 'Three Cry toxins in two types of *Bacillus thuringensis* strain M109 perfentially kill human hepatocyte and uterus cervix cancer cells', *Bioscience, Biotechnology and Biochemistry Journal,* 74, pp. 494-498.

Neher, E. and Sakmann, B. (1976) 'Single-channel currents recorded from membrane of denervated frog muscle fibres', *Nature,* 260(5554), pp. 799-802.

Ohba, Mizuki and Uemori (2009) 'Parasporin, a new anticancer protein group from *Bacillus thuringiensis*', *Anticancer research,* 29(1), pp. 427-434.

Okumura, S., Koga, H., Inouye, K. and Mizuki, E. (2014) 'Toxicity of parasporin-4 and health effects of pro-parasporin-4 diet in mice', *Toxins (Basel),* 6(7), pp. 2115-26.

Okumura, S., Ohba, M., Mizuki, E., Crickmore, N., Cote, J. C., Nagamatsu, Y., Kitada, S., Sakai, H., Harata, K. and Shin, T. (2010) *Parasporin nomenclature* (Accessed: 15 December 2018 2018).

Okumura, S., Saitoh, H., Ishikawa, T., Inouye, K. and Mizuki, E. (2011) 'Mode of action of parasporin-4, a cytocidal protein from *Bacillus thuringiensis*', *Biochimica et biophysica acta (BBA),* 1808(6), pp. 1476-82.

Okumura, S., Saitoh, H., Ishikawa, T., Mizuki, E. and Inouye, K. (2008) 'Identification and characterization of a novel cytotoxic protein, parasporin-4, produced by *Bacillus thuringiensis* A1470 strain', *Biotechnology annu Review,* 14, pp. 225-52.

Pacheco, S., Gomez, I., Arenas, I., Saab-Rincon, G., Rodriguez-Almazan, C., Gill, S. S., Bravo, A. and Soberon, M. (2009) 'Domain II loop 3 of *Bacillus thuringiensis* Cry1Ab toxin is involved in a "ping pong" binding mechanism with *Manduca sexta* aminopeptidase-N and cadherin receptors', *Journal of biological chemistry,* 284(47), pp. 32750-32757.

Palma, L. and Berry, C. (2016) 'Understanding the structure and function of *Bacillus thuringiensis* toxins', *Toxicon,* 109, pp. 1-3.

Palma, L., Muñoz, D., Berry, C., Murillo, J. and Caballero, P. (2014) '*Bacillus thuringiensis* toxins: an overview of their biocidal activity', *Toxins,* 6(12), pp. 3296-3325.

Pardo-Lopez, L., Munoz-Garay, C., Porta, H., Rodriguez-Almazan, C., Soberon, M. and Bravo, A. (2009) 'Strategies to improve the insecticidal activity of Cry toxins from *Bacillus thuringiensis*', *Peptides,* 30(3), pp. 589-595.

Pardo-Lopez, L., Soberon, M. and Bravo, A. (2013) '*Bacillus thuringiensis* insecticidal three-domain Cry toxins: mode of action, insect resistance and consequences for crop protection', *Fems microbiology reviews,* 37(1), pp. 3-22.

Pigott, C. R. and Ellar, D. J. (2007) 'Role of receptors in *Bacillus thuringiensis* crystal toxin activity', *Microbiology and molecular biology reviews,* 71(2), pp. 255-265.

Pigott, C. R., King, M. S. and Ellar, D. J. (2008) 'Investigating the properties of B*acillus thuringiensis* Cry Proteins with novel loop replacements created using combinatorial molecular biology', *Applied and Environmental Microbiology,* 74(11), pp. 3497.

Poornima, Selvanayagam and Shenbagarathai (2010) 'Identification of native *Bacillus thuringiensis* strain form south India having specific cytocidal activity aginst cancer cells', *Journal of Microbiology,* 109, pp. 348-354.

Porta, H., Cancino-Rodezno, A., Soberón, M. and Bravo, A. (2011) 'Role of MAPK p38 in the cellular responses to pore-forming toxins', *Invertebrate Neuropeptides XI,* 32(3), pp. 601-606.

Rajamohan, F., Alzate, O., Cotrill, J. A., Curtiss, A. and Dean, D. H. (1996a) 'Protein engineering of *Bacillus thuringiensis* delta-endotoxin: mutations at domain II of CryIAb enhance receptor affinity and toxicity toward gypsy moth larvae', *Proceedings of the National Academy of Sciences of the United States of America,* 93(25), pp. 14338-14343.

Rajamohan, F., Cotrill, J. A., Gould, F. and Dean, D. H. (1996b) 'Role of domain II, loop 2 residues of *Bacillus thuringiensis* CryIAb delta-endotoxin in reversible and irreversible binding to *Manduca sexta* and Heliothis virescens', *The Journal of Biological Chemistry,* 271(5), pp. 2390-2396.

Rajamohan, F., Hussain, S. R. A., Cotrill, J. A., Gould, F. and Dean, D. H. (1996c) 'Mutations at domain II, loop 3, of *Bacillus thuringiensis* CryIAa and CryIAb delta-endotoxins suggest loop 3 is involved in initial binding to lepidopteran midguts', *Journal of Biological Chemistry,* 271(41), pp. 25220-25226.

Rang, C., Vachon, V., de Maagd, R. A., Villalon, M., Schwartz, J. L., Bosch, D., Frutos, R. and Laprade, R. (1999) 'Interaction between functional domains of *Bacillus thuringiensis* insecticidal crystal proteins', *Applied environmental microbiology,* 65(7), pp. 2918-25.

Roh, J. Y., Nair, M. S., Liu, X. Y. S. and Dean, D. H. (2009) 'Mutagenic analysis of putative domain II and surface residues in mosquitocidal *Bacillus thuringiensis* Cry19Aa toxin', *Fems Microbiology letters,* 295(2), pp. 156-163.

Rother, K., Rother, M., Boniecki, M., Puton, T. and Bujnicki, J. M. (2011) 'RNA and protein 3D structure modeling: similarities and differences', *Journal of molecular modeling*, 17(9), pp. 2325-2336.

Rubio-Infante, N. and Moreno-Fierros, L. (2016) 'An overview of the safety and biological effects of *Bacillus thuringiensis* Cry toxins in mammals', *Journal of Applied Toxicology,* 36(5), pp. 630-648.

Saitoh, Okumura, Ishikawa, Akao, Mizuki and Ohba (2006) 'Investigation of a novel B*acillus thurigensis* gene encoding a parasporal protein, parasporin-4, that preferentially kills human leukemic T Cells', *Bioscience, Biotechnology and biochemistry Journal,* 70(12), pp. 2935-2941.

Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D. R. and Dean, D. H. (1998) '*Bacillus thuringiensis* and its pesticidal crystal proteins', *Microbiology and molecular biology reviews: MMBR,* 62(3), pp. 775-806.

Schwede, T., Diemand, A., Guex, N. and Peitsch, M. C. (2000) 'Protein structure computing in the genomic era', *Research in Microbiology*, 151(2), pp. 107-112.

Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) 'SWISS-MODEL: an automated protein homology-modeling server', *Nucleic acids research,* 31(13), pp. 3381-5.

Sher, D., Fishman, Y., Zhang, M., Lebendiker, M., Gaathon, A., Mancheño, J. and Zlotkin, E. (2005) *Hydralysins, a new category of -pore-forming toxins in Cnidaria.*

Shokry, A. M., Ismail, M. A., Yassin, H. M., Mostafa, S. A., Salama, M. S. and Shahin, M. A. (2018) 'Bioinformatics analysis using homology modeling to predict the three dimensional structure of *Spodoptera littoralis* (*Lepidoptera:Noctuidae*) Aminopeptidase N Receptor'. *Egyptian Journal of Genetics and Cytology*, 41,pp. 269-284

Soberón, M., Fernández, L. E., Pérez, C., Gill, S. S. and Bravo, A. (2007) 'Mode of action of mosquitocidal *Bacillus thuringiensis* toxins', *Toxicon,* 49(5), pp. 597-600.

Soberon, M., Gill, S. S. and Bravo, A. (2009) 'Signaling versus punching hole: How do *Bacillus thuringiensis* toxins kill insect midgut cells?', *Cell Mol Life Sci,* 66(8), pp. 1337-49.

Soberón, M., López-Díaz, J. A. and Bravo, A. (2013) 'Cyt toxins produced by Bacillus thuringiensis: A protein fold conserved in several pathogenic microorganisms', *Peptides,* 41, pp. 87-93.

Soberón, M., Pardo, L., Munoz-Garay, C., Sanchez, J., Gomez, I., Porta, H. and Bravo, A. (2010) 'Pore formation by Cry toxins', *Advances in Experimental Medicine and Biology,* 677, pp. 127-42.

Soberón, M., Portugal, L., Garcia-Gómez, B.-I., Sánchez, J., Onofre, J., Gómez, I., Pacheco, S. and Bravo, A. (2018) 'Cell lines as models for the study of Cry toxins

from *Bacillus thuringiensis*', *Insect Biochemistry and Molecular Biology*, 93, pp. 66-78.

Souissi, W. 2018. Cytocidal activity of Cry41Aa, an anticancer toxin from Bacillus thuringiensis. Doctoral, University of Sussex

Then, C. (2010) 'Risk assessment of toxins derived from Bacillus thuringiensis-synergism, efficacy, and selectivity', *Environmental Science and Pollution Research,* 17(3), pp. 791-797.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) 'CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice', *Nucleic acids research,* 22(22), pp. 4673-4680.

Torres, J. B., Ruberson, J. R. and Adang, M. J. (2006) 'Expression of *Bacillus thuringiensis* Cry1Ac protein in cotton plants, acquisition by pests and predators: a tritrophic analysis', *Agricultural and Forest Entomology,* 8(3), pp. 191-202.

Uemori, Maeda, Yasutake, Ohgushi, Kagoshima, Mizuki and Ohba 'Characterization of Cancer Cell-Killing Activity Associated with Parasporal Proteins of Novel *Bacillus thuringiensis* Isolates', *6th Pacific Rim Conference on the Biotechnology of Bacillus thuringiensis and its Environmental Impact*, Victoria BC, 122-123.

Vachon, V., Laprade, R. and Schwartz, J.-L. (2012a) 'Current models of the mode of action of *Bacillus thuringiensis* insecticidal crystal proteins: A critical review', *Journal of Invertebrate Pathology,* 111(1), pp. 1-12.

Vachon, V., Laprade, R. and Schwartz, J. L. (2012b) 'Current models of the mode of action of *Bacillus thuringiensis* insecticidal crystal proteins: A critical review', *Journal of Invertebrate Pathology,* 111(1), pp. 1-12.

Warren, G. W., Koziel, M. G., Mullins, M. A., Nye, G. J., Carr, B., Desai, N. M., Kostichka, K., Duck, N. B., Estruch, J. J. and Corp, N. (1998) 'Auxiliary proteins for enhancing the insecticidal activity of pesticidal proteins'. U.S. Patent 5,pp. 696–770

Wolfersberger, M., Luethy, P., Maurer, A., Parenti, P., Sacchi, F. V., Giordana, B. and Hanozet, G. M. (1987) 'Preparation and partial characterization of amino acid transporting brush border membrane vesicles from the larval midgut of the cabbage butterfly (*Pieris brassicae*)', *Comparative biochemistry and physiology part A: Physiology,* 86(2), pp. 301-308.

Wolfersberger, M. G. (1984) 'Enzymology of plasma membranes of insects intestinal cells', *Animal zoology,* 24, pp. 187-197.

Wong, R. S. Y., Mohamed, S. M., Nadarajah, V. D. and Tengku, I. A. T. (2010) 'Characterisation of the binding properties of *Bacillus thuringiensis* 18 toxin on leukaemic cells', *Journal of experimental and clinical cancer research,* 29, pp. 11.

Wu, J.-Y., Zhao, F.-Q., Bai, J., Deng, G., Qin, S. and Bao, Q.-Y. (2007) 'Adaptive evolution of Cry genes in *Bacillus thuringiensis*: implications for their specificity determination', *Genomics, proteomics and bioinformatics,* 5(2), pp. 102-110.

Wu, S.-J. and Dean, D. H. (1996) 'Functional significance of loops in the receptor binding domain of *Bacillus thuringiensis* CryIIIA δ-endotoxin', *Journal of molecular biology,* 255(4), pp. 628-640.

Wu, S.-J., Koller, C. N., Miller, D. L., Bauer, L. S. and Dean, D. H. (2000) 'Enhanced toxicity of *Bacillus thuringiensis* Cry3A δ-endotoxin in *coleopterans* by mutagenesis in a receptor binding loop', *FEBS Letters,* 473(2), pp. 227-232.

Xie, R. Y., Zhuang, M. B., Ross, L. S., Gomez, I., Oltean, D. I., Bravo, A., Soberon, M. and Gill, S. S. (2005) 'Single amino acid mutations in the cadherin receptor from *Heliothis virescens* affect its toxin binding ability to Cry1A toxins', *Journal of biological chemistry,* 280(9), pp. 8416-8425.

Xu, C., Wang, B.-C., Yu, Z. and Sun, M. (2014) 'Structural insights into *Bacillus thuringiensis* Cry, Cyt and parasporin toxins', *Toxins,* 6(9), pp. 2732-2770.

Yamashita 04/06/2009 2009. *RE: about expression of parasporin (PS3).* Type to Crickmore, N.email.

Yamashita, S. (2005) 'Typical three-domain Cry proteins of *Bacillus thuringiensis* strain A1462 exhibit cytocidal activity on limited human cancer cells', *Journal of biochemistry,* 138(6), pp. 663-672.

Yamashita, S., Akao, T., Mizuki, E., Saitoh, H., Higuchi, K., Park, Y. S., Kim, H. S. and Ohba, M. (2000) 'Characterization of the anti-cancer-cell parasporal proteins of a *Bacillus thuringiensis* isolate', *Canadian journal of microbiology,* 46(10), pp. 913-919.

Yasutake, Binh, Kagoshima, Uemori, Ohgushi, Maeda, Mizuki, Yu, M. and Ohba 'Occurrence of *Bacillus thuringiensis* producing parasporin, cancer cell-killing Cry proteins, in Vietnam', *6th Pacific Rim conference on the biotechnology of Bacillus thuringiensis and its environmental impact*, Victoria BC, 124-125.

Zalunin , A., Revina, P., Kostina, I. and Chestukhina, G. (2004) 'Peculiarities of Cry proteins to be taken into account during their *in vivo* and *in vitro* study', *International organization for biological and integrated control of noxious animal and plants, West Palearctic Regional Section (IOBC/WPRS),* 27(3),pp 177-186 .

Zarubin, T. and Han, J. (2005) 'Activation and signaling of the p38 MAP kinase pathway', *Cell Research,* 15, pp. 11.

Zhang, X., Candas, M., Griko, N. B., Rose-Young, L. and Bulla Jr, L. A. (2005) 'Cytotoxicity of *Bacillus thuringiensis* Cry1Ab toxin depends on specific binding of the toxin to the cadherin receptor BT-R1 expressed in insect cells', *Cell death and differentiation,* 12, pp. 1407.

Zhang, X., Candas, M., Griko, N. B., Taussig, R. and Bulla, L. A. (2006) 'A mechanism of cell death involving an adenylyl cyclase/PKA signaling pathway is induced by the Cry1Ab toxin of *Bacillus thuringiensis'*, *Journal of proceedings of the national academy of sciences of the United States of America*, 103(26), pp. 9897-902.

Zhou, Z., Liu, Y., Liang, G., Huang, Y., Bravo, A., Soberón, M., Song, F., Zhou, X. and Zhang, J. (2017) 'Insecticidal Specificity of Cry1Ah to *Helicoverpa armigera* is determined by binding of APN1 via domain II loops 2 and 3', *Applied Environmental Microbiology,* 83(4).