Title
Genome Sequence and the Identification of Mosquitocidal Toxin Operons in Clostridium bifermentans subsp. malaysia

Permalink
https://escholarship.org/uc/item/4w52t7jq

Author
Chawla, Swati

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Genome Sequence and the Identification of Mosquitocidal Toxin Operons in *Clostridium bifermentans* subsp. *malaysia*

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Environmental Toxicology

by

Swati Chawla

March 2015

Dissertation Committee:
    Dr. Sarjeet S. Gill, Chairperson
    Dr. Michael E. Adams
    Dr. Thomas Girke

The Dissertation of Swati Chawla is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ABSTRACT OF THE DISSERTATION

Genome Sequence and the Identification of Mosquitocidal Toxin Operons in *Clostridium bifermentans* subsp. *malaysia*

by

Swati Chawla

Doctor of Philosophy, Graduate Program in Environmental Toxicology
University of California, Riverside, March 2015
Dr. Sarjeet S. Gill, Chairperson

The management and control of mosquito vectors of human disease currently rely primarily on chemical insecticides. However, larvicidal treatments can be effective, and if based on biological insecticides, they can also ameliorate the risk posed to human health by chemical insecticides. The aerobic bacteria *Bacillus thuringiensis* and *Lysinibacillus sphaericus* have been used for vector control for a number of decades. But development of mosquito resistance to individual toxins has been seen in laboratory and in field. More recently, the anaerobic bacterium *Clostridium bifermentans* subsp. *malaysia* (Cbm) has been reported to have high mosquitocidal activity, and a number of proteins were identified as potentially mosquitocidal. However, the cloned proteins showed no mosquitocidal activity.

I determined genome sequence of the only known anaerobic mosquitocidal bacterium, Cbm by the next generation sequencing approach. I showed that the genome consists of a 3.7-Mb chromosome, seven small plasmids and a 109-kb megaplasmid that contains the two toxin operons responsible for mosquitocidal activity. Also, additional virulence-related factors could be identified for further work.

I co-authored and published that four toxins encoded by the Cry operon, *Cry16A*, *Cry17A*, *Cbm17.1*, and *Cbm17.2*, are all required for toxicity, and these toxins collectively show remarkable selectivity for *Aedes* rather than *Anopheles* mosquitoes, even though Cbm is more toxic to *Anopheles*. I also tried to improve the mosquitocidal activity of Cbm Cry operon by protein engineering of surface loop residues in domain II of Cry17Aa and wanted to test on 4 different species of important human disease vectors, *Aedes aegypti*, *Anopheles gambiae, Anopheles stephensi* and *Culex quinquefasciatus*.

I also identified the toxins that could target *Anopheles* are different from those expressed by the Cry operon and investigated the possibility that the genes encoded in second operon on pClosMP, the Clostridial Mosquitocidal Protein (Cmp) operon, could be toxic to *Anopheles*. The Cmp operon encodes seven genes, including a 150-kDa mosquitocidal toxin (*CMP1*) and *HA* gene. I show the three genes, *CMP1, CNTNH,* and *CHA* that play an important role in botulinum neurotoxicity are not active individually against mosquitoes. We need to express the native full length 16kb Cmp operon and CMP1-CNTNH-HA constructs to answer this question.

In summary, based on data in the literature and my work, I demonstrate that Cbm harbors its mosquitocidal activity on a large 109kb plasmid that encodes for two toxin operons. The Cry operon toxins work as a complex and are active against *Aedes* larvae. The Cmp operon toxins activity against *Aedes* and *Anopheles* larvae remains a vast field of study to be further explored.

Table of Contents

List of Figures

List of Tables

General Introduction

1. Introduction to Mosquito-borne diseases

Mosquito-borne diseases significantly impacted human civilization long before mosquitoes were known to be efficient vectors of disease. Diseases such as dengue and malaria still remain serious public health problems in the world. Yellow fever, filariasis, chikungunya, West Nile fever, and leishmaniasis are just a few vector-borne diseases that have been responsible for significant human disease and death (Gubler, 1991). Even with sizeable medical and technological advances, emerging and re-emerging mosquito-borne diseases continue to cause mortality and significantly impact human populations.

Mosquito-borne parasites have had a considerable resurgence since the 1970s. Malaria is the most important vector-borne disease due to its global distribution and the number of people affected worldwide. There were an estimated 207 million cases of malaria worldwide in 2012, with and 627,000 malaria deaths worldwide. The 1.2 billion at high risk (>1 case per 1000 population) are living mostly in the African Region (47%) and the South-East Asia Region (37%). (Figure 1.1; WHO, 2012). Malaria parasites are transmitted to humans by mosquito species of the genus *Anopheles*. At least 40 anopheline species are important human malaria vectors worldwide. Although malaria is a preventable and curable disease, it remains one of the most prevalent parasitic diseases of humans (Gardiner *et al.,* 2005).

Mosquito-borne viruses as a group, however, are as important, if not more so than malaria in terms of numbers of infections, varieties of vector species and global distribution. But most importantly for arboviruses, particularly dengue, yellow fever and chikungunya viruses are all spread by *Aedes aegypti*, a mosquito that lives in close association with humans in urban and sub-urban environments. Therefore, these diseases are generally considered diseases of urban areas, and its epidemiology is highly related to the biology of the mosquito vector, the environment, and human behavior. Yellow fever, for example, is a serious disease in Africa and South America: 200,000 infections annually resulting in 30,000 deaths in spite of vaccine usage (WHO, 1998). Dengue is the second most important tropical disease that causes more illness and death than any other arbovirus (WHO, 2002). The World Health Organization estimates that 50–100 million dengue infections occur each year and that almost half the world's population lives in countries where dengue is endemic. While dengue is a global concern, with a steady increase in the number of countries reporting the disease, currently close to 75% of the global population exposed to dengue are in the Asia-Pacific region (Figure 1.2, WHO, 2012). There is no effective vaccine for dengue fever and the control of their vector, *Ae. aegypti*, is the only reasonable preventive option. However, even though there is a highly effective and safe yellow fever vaccine available since 1930s, present day epidemics still occur in parts of Africa and South America incurring significant mortality (Barrett and Higgs, 2007).

In the last decade, West Nile virus has made resurgence in Europe and has emerged in North America (Lanciotti *et al.,* 1999; Nash *et al.,* 2001). West Nile virus, an

arbovirus, is spread by the mosquito species of the genus *Culex* and has traditionally only caused severe disease in the elderly, the young, and the immunocompromised.

Numerous factors have contributed to the re-emergence of vector-borne diseases. Increases in insecticide and drug resistance, virus evolution, and alteration in global climate have all directly impacted transmission of vector-borne diseases (Porter et al., 1993; Gubler, 1997). Additional problems are lack of a pro-active, prevention-based public health policy with substantial infrastructure, as well as the decline of effective mosquito control programs in the past 40 years (Gubler, 2004). Global demographic and societal changes have also significantly affected emergence and resurgence of mosquito-borne diseases. Unprecedented global population growth, along with unplanned and uncontrolled urbanization in tropical developing countries, has created ideal conditions for increased vector densities and vector-borne disease transmission potential (WHO, 2012). Relative ease of global travel has also contributed to the transmission potential of vectors and disease pathogens.

Consequently, many attempts to control mosquito abundance have been made using various methods. Early methods for controlling the spread of mosquito-borne diseases relied primarily upon synthetic chemical insecticides. These were widely deployed with excellent results, but their use ultimately caused adverse environmental effects and the evolution of insecticide resistance in many targeted mosquito species (Wirth et al., 1998, 2010; Yang et al., 2007). The need to control mosquito populations continues, however the insecticides that can be safely used in such efforts are extremely limited. Because of this, interest has grown in the bacterial insecticides, primarily

Bacillus thuringiensis subsp. *israelensis* (Bti) and *Lysinibacillus sphaericus* (Ls) that produce crystalline proteins that are toxic against various species of *Culex, Anopheles* and *Aedes* larvae and are safe for the environment.

2. *Bacillus thuringiensis* Insecticidal Toxins

2.1 *Bacillus thuringiensis*

Bacillus thuringiensis (Bt) is a Gram-positive bacterium, which was discovered and isolated by Ishiwata Shigetane in 1901 as the cause of the disease that was killing the silkworm populations (Luthy *et al.,* 1981). This bacterium was named *B. thuringiensis* by Berliner in 1911 and classified as a member of the *Bacillus cereus* group in the Bacillaceae family, which contains among others, *B. anthracis* and *B. cereus* (Rasko et al., 2005). Bt is pathogenic to insects. In adverse conditions, this bacterium forms endospores and, at the same time, produces crystalline parasporal inclusions that contain one or more proteins toxic by ingestion for insects (Schnepf et al, 1998). In 1956, the crystal protein inclusions that are produced during sporulation were associated with the insecticidal activity (Angus, 1956). The expression of these inclusion proteins was shown to be encoded by transmissible plasmids (Gonzalez et al., 1982; Berry *et al.,* 2002). These inclusions contain one or more proteins called Cry and Cyt toxins (Hofte *et al.,* 1989). These proteins are highly selective to the target insect, but are harmless to humans and vertebrates such that they have been used to control lepidopteran, dipteran and coleopteran insect pests in agriculture and public health (Crickmore *et al.,* 1998). Among them, *Bacillus thuringiensis* subsp. *israelensis* (Bti) was isolated by Goldberg and

Margalit in 1977 (Goldberg et al., 1977) and has high toxicity to Dipteran insects, primarily mosquito and black fly larvae (Mittal, 2003; Lacey, 2007). The success of Bt as bioinsecticide came with the development of Bt-crops that express the cry gene resulting in crops that resist insect attack including borers that were difficult to control with topical Bt-formulations leading to the commercial release of Bt-crops in 1995 (Tabashnik, et al., 2008; Bravo et al., 2013). The first-generation Bt-crops target lepidopteran pests and produce only one Bt toxin, for example, cotton producing Bt toxin Cry1Ac and corn producing Bt toxin Cry1Ab. Some newer varieties of Bt crops produce two Bt toxins, the list of which is available at http://www.epa.gov/pesticides/biopesticides/pips/pip_list.htm.

2.2 *Bacillus thuringiensis* **Toxin Classification**

The insecticidal proteins in the crystalline bodies produced during sporulation have been shown to contain two types of insecticidal proteins namely Cry toxins and Cyt-toxins. Schnepf and Whiteley in1981, confirmed that the insecticidal ability of Bt by cloning and expressing the first toxin gene from *Bacillus thuringiensis* var. *kurstali* HD-1 in *E. coli* which showed insecticidal activity to *Manduca sexta*. Since this discovery, a great number of other genes have cloned and expressed and the process of Bt toxin gene discovery is still ongoing. This created a need to organize the ever growing data. After many attempts to produce an organized robust systematic nomenclature of Bt insecticidal genes, Crickmore et al. in 1998 came up with a system that is based on amino acid sequence. Cry toxins are classified by their primary amino acid sequence and more than 700 different cry gene sequences have been classified into 70 groups (Crickmore et al.,

2011) where toxins belonging to each Cry group share less than 40% amino acid identity

with proteins from other groups (Crickmore et al., 1998). Within each group, a capital

letter (Cry1A, Cry1B etc) is given when they share less than 70% identity. A small letter

(Cry1Aa, Cry1Ab etc) is given when toxins share more than 70% but less than 95%

identity. Phylogenetic analysis of Cry protein sequences showed that the whole family of

Cry proteins belong to four non-phylogenetically related protein families, the family of

three domain Cry toxins (3D), the family of mosquitocidal Cry toxins (Mtx), the family

of the binary-like (Bin) and the Cyt family of toxins (Bravo et al., 2005). Some Bt strains

produce additional insecticidal toxins named VIP during vegetative growth, and as these

proteins do not form parasporal crystals, they were not named Cry toxins. Three VIP

toxins have been characterized as VIP1/VIP2, which together compose a binary toxin,

and VIP3 (Estruch et al., 1996).

The binary toxins and related toxins Mtx1, Mtx2, Mtx3 toxins (de Maagd et al.,

2003; Palma et al., 2014) produced by *B. sphaericus* form crystals and do not have a

tertiary structure in three domains, but both proteins are required to produce the

insecticidal effect. They are active against mosquitoes and show homology to pore

forming toxins such as the aerolysin from *Aeromonas hydrophila* and epsilon toxin of

Clostridium perfringens. For example, Cry34/35ab1 toxins that shows toxicity towards

western corn rootworm, *Diabrotica virgifera virgifera* LeConte. Cry34 does not show

homology with any other Cry protein but Cry35 shows homology with other Bin toxins

(Ellis et al., 2002) and shares with them a similar three-dimensional folding structure (de

Maagd et al., 2003; Palma et al., 2014). Both Cry34 and Cry35 are required for activity.

6

A website has been established and is frequently updated as new genes that are discovered (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/).

2.3 Structure of *B. thuringiensis* Insecticidal Toxins

Cry toxins are a family of related proteins that are synthesized as 130 – 140 kDa protoxins and are activated by removal of the C-terminal half of the protein in the insect gut, leaving an active toxin of about 60kDa (Hofte and Whiteley, 1989). Alignment of the sequences of Cry toxins of different specificities show five highly conserved sequence blocks in the active N-terminal half of the protein, separated by hyprvariable sequences. This active N-terminal half of the protein shows a three domain structure as resolved by X-ray crystallographic methods. Interestingly, the C-terminal half of the protein that is cleaved in the insect gut shows very high conservation of residues. Hence, Hofte and Whiteley in 1989 proposed that the C-terminal half may be intimately involved in crystal formation, and the conservation of sequences in this part of the molecule is apparently sufficient to allow the coassembly of different Cry proteins in the same crystals.

The first 3-D crystal structure of Cry-proteins reported was for the coleopteran specific Cry3Aa (Li et al., 1991). Currently the 3-D crystal structures of coleopteran specific Cry3Bb (Galitsky et al., 2001) and Cry8Ea1 (Guo et al., 2009) (Figure 1.3); dipteran specific Cry4Ba (Boonserm et al., 2005) (Figure 1.4) and Cry4Aa (Boonserm et al., 2006); lepidopteran specific Cry1Aa (Grochulski et al., 1995); and lepidopteran/dipteran specific Cry2Aa (Morse et al., 2001) have been resolved through X-ray crystallographic methods of their activated forms. The structures of Cyt toxins,

Cyt2Aa toxin (Li et al., 1996) and Cyt1Aa (Cohen et al., 2011)were also determined by

X-ray crystallography. Though different Cry toxins have been shown to have specific

targets in their insecticidal activity, the overall 3-D fold of many of them has been shown

to be the same comprising of three domains (de Maagd et al., 2003).

1. Domain I is compose of seven α-helices in which the central helix-α5 is
   hydrophobic and is encircled by six other amphipathic helices. This domain has
   been shown to be involved in toxin oligomerization, membrane insertion and pore
   formation (Schnepf et al., 1998).

2. Domain II is made up of three antiparallel β- sheets packed together (Li et al.,
   1991). Two of the sheets are composed of four strands in a Greek key motif and
   are solvent exposed (Boonserm et al., 2006). The third sheet packs against domain
   I and is arranged in a Greek-key-like motif with three strands and a short alpha-
   helix (Pigott et al., 2007). The exposed loop regions have been shown to be
   involved in binding to specific larval midgut proteins.

3. Domain III has been shown to contain two antiparallel β-sheets (Boonserm et al.,
   2006). Both sheets are composed of five strands, with the outer sheet facing the
   solvent and the inner sheet packing against domain II. Domain III shows less
   structural variability than domain II, and the main differences are found in the
   lengths, orientations, and sequences of the loops (Boonserm et al., 2005). It has
   been seen that Domains II and III have functions in specific recognition and
   binding to receptors and for domain III, in modulating ion channel activity
   (Schnepf et al., 1998).

2.4 Mechanisms of action of *Bacillus thuringiensis* toxins

The mode of action of Bti is accepted to be a multistage process (Ramírez-Lepe, et al., 2012). The toxic properties of Bti originate in crystalline inclusions produced during sporulation. The crystals and their subunits are inert protoxins and are not biologically active. However, they become activated when ingested by mosquito larvae that die within a few hours of ingestion.

The "classical" mode of action of Cry toxins can be outlined as follows: (i) ingestion of inactive crystal proteins by the larvae; (ii) solubilization of the crystals in the alkaline (pH>10) midgut to release protoxin; (iii) proteolytic cleavage of the insecticidal 130-140kDa protoxin by midgut proteases yielding an activated protease resistant core of about 60-70kDa biologically active toxin (Bravo et al.,2007) and comprised by the three-dimensional structure (de Maagd et al., 2001); (iv) activated toxin binds to specific receptors located on midgut cell membranes (Bravo et al., 2011); (v) a change in the toxin conformation allowing toxin insertion into the midgut epithelium columnar cells (Gomez et al., 2007); (vi) electrophysiological and biochemical evidence suggest that the toxins generate lytic pores in the in microvilli of apical membranes (Bravo et al., 2007), thus disturbing the osmotic balance and causing the cells to swell and lyse; (vii) disruption of the midgut epithelium releases the cell content providing *Bacillus thuringiensis* spores a germinating medium leading to a severe septicemia and insect death (Bravo et al., 2007).

Two models of the mode of action of solubilized and activated Cry toxin have been proposed in Lepidoptera (Figure1.5). The first proposed mechanism is the

sequential binding model detailing the hypothetical mechanism of pore-formation (Soberon et al., 2009), and the second proposed model is the signaling cascade model, which contends that cytotoxicity is mediated by the specific binding of Bt toxins to their cadherin receptors which activate adenylyl cyclase/protein kinase A (Zhang et al., 2006) signaling pathways that lead to necrotic cell death.

Many proteins receptors have been identified, that are potentially able to bind Cry toxins, but the better characterized and validated as functional receptors in different insects are the aminopeptidase N (APN), cadherins, alkaline phosphatase (ALP), glycolipids and recently also an ABC transporter. Cadherin's are a superfamily of glycosylated proteins having repetitions of calcium binding domains, called cadherin repeats. In mosquitoes, our lab has shown that *Aedes aegypti* cadherin is able to bind Cry11Aa and Cry11Ba toxins (Chen et al., 2009a; Likitvivatanavong et al., 2011) and acts as a mediator of in-vivo toxicity for Cry11Aa by using silenced *Aedes* cadherin expression in transgenic mosquitoes (Lee et al., 2014). Our lab also showed that *Aedes* cadherin may mediate Cry4Aa and Cry11Ba toxicity, but is not the main receptor of Cry4Aa, Cry4Ba and Cry11Ba toxin in *Ae. aegypti*. In *Anopheles gambiae*, cadherin is also shown to bind Cry4Ba (Hua et al., 2008). APN is zinc dependent peptidase anchored to the outside of the plasma membrane by a glycosyl phosphatidylinositol (GPI). APN has been shown to bind Cry11Ba toxin in *Anopheles quadrimaculatus* (Abdullah et al., 2006) and in *A. gambiae* (Zhang et al., 2008b). In *Ae. aegypti* it has been shown that APN exhibits binding capacity to Cry11A (Chen et al., 2009b), Cry4A (Bayyareddy et al., 2009) and Cry4Ba (Saengwiman et al., 2011). Like APN, ALP is Also a GPI anchored

metalloenzyme and shows ability to bind Cry toxins. In *Ae. aegypti*, ALP's have been

shown to bind Cry11Aa (Fernández et al., 2006) and Cry4Ba (Dechklar et al., 2011).

ABC (ATP-binding cassette) transporters are proteins anchored to the membrane through

12 transmembrane alpha helices. They use the energy of ATP hydrolysis to export

compounds from the cytoplasm to the cell outside or within intracellular compartments

(Heckel, 2012). McNall and Adang (2003) identified an ABCC2 transporter in *M.sexta* as

possible Cry1Ac-binding molecule. Two field evolved Cry1Ac resistant colonies of

Diamondback moth (*Plutella xylostella*) and the cabbage looper (*Trichoplusia ni*) and a

laboratory selected tobacco budworm (*Heliothis virescens*) Cry1Ac resistant colony were

recently shown to be linked to mutations in the ABCC2 transporter (Baxter et al., 2011;

Gahan et al., 2010).

2.5 Other Mosquitocidal Bacterial Strains and Bti Resistance

Biological control of mosquitoes and blackflies, that transmit severe human

diseases, have relied heavily on the use of *Bacillus thuringiensis* subsp. *israelensis* and *B.

sphaericus* as larvicidal agents (Nicolas et al, 1990). As the genomes of both the bacteria

have been sequenced, the genes encoding dipteran active toxins are conserved and share

great homology (Höfte and Whiteley, 1989). These proteins are highly selective to the

target insect, but are harmless to humans and vertebrates such that they have been used to

control lepidopteran, dipteran and coleopteran insect pests in agriculture and public

health (Crickmore *et al.,* 1998). Some Bt strains, including Bt serovars *israelensis*,

jegathesan, *medellin* and *morrisoni*, were found to be highly toxic to mosquito larvae and

11

contained mosquitocidal crystal proteins. Among them, *Bacillus thuringiensis* subsp. *israelensis* (Bti) is highly toxic to Dipteran insects, primarily mosquito and black fly larvae (Mittal, 2003; Lacey, 2007). *Lysinibacillus sphaericus* has been used successfully in the control of *Culex quinquefasciatus* and *Culex pipiens*. However, resistance to Ls toxins has been observed in the field in a number of countries (Nielsen-Leroux *et al.,* 1995; Rao *et al.,* 1995; Charles and Nielsen-LeRoux, 2000). Bti is primarily used to control A. *aegypti* but is also useful against *Culex* and *Anopheles* mosquitoes (Porter, 1993) (Figure 1.6). The primary threat to the use of Cry toxins in transgenic plants is the appearance of insect resistance. Evolution of resistance to Bt-crops in the field has been documented for at least five different insect species, *Busseola fusca* (van Rensburg, et al., 2007), *Spodoptera frugiperda* (Storer et al., 2010) and *Diabrotica virgifera virgifera* (Gassmann et al., 2011) to Bt-transgenic maize, and *Helicoverpa zea* (Tabashnik et al., 2008) and *Pectinophora gossypiella* (Bagla, 2010) to Bt-transgenic cotton. Therefore, an alternative for the in vitro genetic evolution of Cry toxins with the aim of enhancing toxicity against specific pests, to kill novel targets or to recover toxicity in the case of the appearance of resistance in the field, is screening and isolation of novel toxin protein in nature (Pardo-López et al., 2009). Furthermore, the frightening emergence of drug-resistant malaria parasites and insecticide-resistant mosquitoes and the lack of a malaria and dengue vaccine, together with the toxic effects of chemicals insecticides to non-target flora and fauna, have brought the isolation of novel mosquitocidal bacterial control firmly back into the conversation.

There are other entomopathogenic bacteria like *Bacillus cereus* that do not form crystals, but their spores can cause septicemia in insects (Helgason et al., 2000). Various *Paenibacillus* spp. also produce antimicrobial substances that affect a wide spectrum of micro-organisms (Giardin et al., 2002; Piuri et al., 1998) such as fungi, soil bacteria, plant pathogenic bacteria and even important anaerobic pathogens as *Clostridium botulinum*. Other biological insecticide products that are commercially available are based on *Serratia entomophila*, *Xenorhabdus* and *Photorhabdus* spp. associated with entomopathogenic nematodes also produce potent insecticidal toxins that could represent additional alternatives for insect control (Ffrench-Constant, et al., 2000). Recently, two strains of *Clostridium bifermentans* toxic to mosquito larvae were isolated from a soil sample collected from secondary forest floor. One strain was designated as *Clostridium bifermentans* serovar *malaysia* (Cbm) and the other *Clostridium bifermentans* serovar *paraiba* (Cbp) according to its specific H antigen (Seleena *et al.,* 1997). These clostridial strains are not toxic to vertebrates (Thiery et al., 1992a) but highly toxic to mosquito larvae. Cbp is toxic to *Anopheles maculatus* whereas its toxicity to *Ae. aegypti* and *C. quinquefasciatus* larvae is 20 and 3 times lower, respectively. Cbm toxins are highly toxic to *Anopheles* and *Aedes* species. Its toxicity against *An. stephensi* is about 10 times higher, but its toxicity against *Culex pipiens* or *Aedes aegypti* is 10 times lower, than that of Bti (Thiery et al., 1992b).

3. Clostridia

The genus *Clostridium* consists of Gram-positive, rod-shaped bacteria in the Phylum Firmicutes. It contains over 100 different species which have been characterized and classified according to their 16s rDNA sequences. All species are strictly anaerobic and form endospores. Clostridia have been one of the most studied bacterial classes since the early identification of its role in a number of human diseases (Melville, 2007).

All pathogenic clostridial species produce protein exotoxins that play an important role in their pathogenicity. Several species are commonly listed as important human pathogens. Of all clostridial species, the 13 that are considered as major pathogens in man and/or animals are: *C. botulinum*, *C. chauvoei*, *C. colinum*, *C. difficile*, *C. haemolyticum*, *C. histolyticum*, *C. novyi*, *C. perfringens*, *C. piliforme*, *C. septicum*, *C. sordellii*, *C. spiroforme*, and *C. tetani*.

3.1 Clostridial Genomes

Clostridial genomes generally consist of a single circular double-stranded chromosome in the size range of 3.5 to 7.0 Mega base pairs (Mbp), often accompanied by plasmids of variable sizes (Katayama et al., 1995; Nölling et al., 2001). As most bacteria, *Clostridium sp.* contains integrated bacteriophages and plasmids in the size range of 3 – 200 kilo base pairs (Kbp). Based on the relative content of G+C of the chromosome, *Clostridium* species can be classified into two groups: the extremely low G+C group (ranging between 24% and 31% G+C contents) includes *C. perfringens*, and *C. acetobutylicum*, and the second group, showing slightly higher G+C contents (ranging

between 32 and 40%) includes thermophiles such as *C. thermocellum* and *C. cellolyticum* (Young et al., 1989). AT-rich stretches seem to cover mainly non-coding regions. The chromosomes contain 3000 – 4000 open reading frames and encode 10 – 14 rRNAs. Due to their exceptionally low G+C contents, Clostridia have particular codon usage. For example in *C. acetobutylicum* and *C. perfringens* arginine is almost always encoded by AGA and CGC, CGA and CGG are not used. Clostridia generally use AUG (Met), GUG (Val) and UUG (Leu) as translational start codons. Particularly striking are the ribosome binding sites (RBS) of Clostridia that contain preferentially G residues; the most redundant RBS encountered being GGAGG and GGGGG. The preferred initiation and termination sequences are AUG and UAA, and the degenerate codons ending in C or G are not used (Thompson et al., 1990). The nucleotide sequence of the neurotoxin gene has been identified for all toxin types, as well as some of the surrounding genes (Binz et al., 1990a and 1990b; Thompson et al., 1990; Poulet et al., 1992; Whelan et al., 1992; Campbell et al., 1993). Currently in NCBI Database, there are sequences of 38 clostridial genomes in various stages of progress and completion.

3.2 *Clostridium botulinum* Toxin

Botulism has been known for hundreds of years and was first described in the early nineteenth century. The illness was then named after the Latin term for sausage (botulus). It is a rare but serious paralytic illness caused by the botulinum neurotoxins (BoNTs), which are produced *by Clostridium botulinum* and other species under anaerobic conditions (Hatheway, 1990). The toxins are large proteins that act on

cholinergic neuromuscular junctions and are responsible for inhibition of the

neurotransmission. Botulism infections are characterized by progressive flaccid paralysis,

firstly of the facial muscle. In most severe cases it can spread towards the limbs and

cause respiratory failure, thereby leading to death.

C. botulinum strains are commonly divided in four different groups (groups I, II,

III, and IV) based on physiologic characteristics (Collins and East, 1998). The toxins

produced are ordered into seven serologically distinct groups (serotypes A to G), based

on recognition by polyclonal serum (Hatheway, 1990). Other Clostridia species have

been shown to produce botulinum neurotoxins, namely *C. butyricum* (BoNT/E), *C.*

baratii (BoNT/F), and *C. argentinense* (BoNT/G) (Collins and East, 1998). Most

outbreaks of human botulism are caused by group I (proteolytic) or II (non-

proteolytic) *C. botulinum*. Group III organisms mainly cause diseases in animals. The

gene coding for BoNT serotypes A, B, E, and F are located on the bacterial chromosome.

Strains producing serotypes C and D have the BoNT gene encoded on a phage genome

(Hill et al., 2007).

3.2.1 Structure of BoNT Toxin

Botulinum toxins are produced in complexes with associated non-toxic proteins

(NAPs), called the progenitor toxin complexes (PTCs). The PTC is composed of a BoNT

molecule and several auxiliary proteins termed neurotoxin-associated proteins (NAPs).

BoNT/A–D and G are produced in bacteria together with four NAPs, which include non-

toxic non-hemagglutinin (NTNHA) protein and three hemagglutinins (HAs: HA33,

HA17 and HA70, also known as HA1, HA2 and HA3, respectively) (Collins and East

1998). In contrast, BoNT/E and F do not have the HA genes. Instead, they contain genes

of OrfX1/X2/X3, whose expression and function are still unknown (Kubota et al., 1998).

The neurotoxin genes (BoNT) and genes encoding for the NAPs (except HAs) in *C.*

botulinum are located in an operon and are expressed together (Couesnon et al., 2006).

The PTC is transcribed as two divergent polycistronic messages, with the NTNH and

BoNT in one and HA70, HA17 and HA33 in the other. A gene called botR, also located

in the locus, has been found in all types except non-proteolytic type E and acts as a

positive regulator (Marvaud et al., 1998). There is great genetic diversity between strains,

even within the same type, and it has been suggested to further divide them into

subgroups (Hill et al., 2007). The seven BoNTs have a high degree of primary sequence

conservation, although all are antigenically distinct (Lacy and Stevens 1999).

The minimally functional PTC is composed of BoNT (~150 kDa) and NTNHA

(~140 kDa). BoNT/A is secreted as a progenitor toxin complex (Figure 1.7 and 1.11) in

one of three sizes, 12S (300 kDa, M-TC), 16S (500 kDa, L-TC), or 19S (900 kDa, LL-

TC), depending on the type and number of NAPs associated with the complex (Oguma et

al., 1999, Tonozuka et al., 2014). The 12S progenitor toxin complex consists of a single

BoNT molecule and one NTNH protein, but lacks the associated proteins responsible for

hemagglutination activity. The 16S progenitor toxin, in addition to the components found

in the 12S complex, contains three hemagglutinin (HA) proteins, HA70, HA33, and

HA17. The secreted and most toxic form 19S complex is believed to be a dimer of two

16S toxins linked by an additional HA33 protein. To summarize, BoNT+NTNH=12S,

12S+HA complex=16S, HAcomplex=6 HA1+3 HA2+3 HA3. Each serotype (A ~ F)

produces toxin complexes as follows: Type A, 12S, 16S, and 19S; types B–D, 12S and

16S; types E and F, 12Sand type G, 16S.

Crystal structures have been reported for full-length BoNT/A (PDB: 3BTA),

BoNT/B (PDB: 1EPW), and BoNT/E (PDB: 3FFZ) (Lacy et al. 1998; Swaminathan and

Eswaramoorthy 2000; Kumaran et al. 2009). The crystal structures of the isolated HAs

have been reported for selected serotypes, including HA-17 (BoNT/D), HA-33 (BoNT/A,

C, and D), and HA-70 (BoNT/C) (Arndt et al. 2005; Hasegawa et al. 2007; Inoue et al.

2003; Nakamura et al. 2009; Nakamura et al. 20011).

3.2.2 Mode of Action of BoNT

BoNT is synthesized as a single polypeptide chain of ~150 kDa and post-

translationally nicked by an unknown protease into a ~50 kDa light chain (LC) and a

~100 kDa heavy chain (HC) (Schiavo, et al., 1992; Rummel and Binz, 2012) when they

are released on cell lysis. All three crystal structures reported are similar in that they

exhibit a modular architecture comprising three domains. The LC is a protease. The HC

is composed of two domains: the N-terminal domain (HN, also known as the

translocation domain) mediates translocation of LC across the endosomal membrane,

whereas the C-terminal domain (HC) is the cell surface receptor binding domain. Primary

sequence analysis of LC revealed an "HELIH+E" signature motif in all BoNTs (Ile is

highly conserved except in BoNT/C and D where it is replaced by Asn and Thr,

respectively). His and the second Glu residues coordinate a Zn2+, while the first Glu

coordinates a water molecule for hydrolysis. This motif is found in a variety of Zn2+-dependent metalloproteases such as thermolysin, which suggests that LCs may utilize a similar enzymatic mechanism (Binz et al. 2002; Kurazono et al. 1992). However, such a motif is not found in NTNHA but is replaced by a highly conserved "KCLIK" motif in NTNHAs (Gu et al. 2012).

Following toxin ingestion, the neurotoxin complex has to evade the acidic conditions in the GI tract, resist proteolytic degradation in the intestines and bind to the epithelium of the intestine to reach the circulatory system and the bind to the neuromuscular neurons to cause a toxic effect. NTNHA directly interacts with BoNT and form an interlocked handshake-like complex, which lends both proteins stability against low pH and digestive proteases (Gu et al., 2012). The three HA NAPs interact with intestinal epithelial cells and play an active role in BoNT transport (Fujinaga et al. 1997; Fujinaga et al. 2004; Niwa et al. 2007). Strong support for the functional role of HAs is provided by the findings that HAs directly bind mucins, epithelium microvilli, oligosaccharides (Fujinaga et al. 2004; Fujinaga et al. 2013; Inoue et al. 2001; Nakamura et al. 2007), Nacetylneuraminic acid (Nakamura et al. 2009), α2-3-sialylated oligosaccharide (Yamashita et al. 2012), galactose (Gal), and N-acetylgalactosamine (Nakamura et al. 2008, 2011) and even gangliosides like GD1, GT1, GQ1depending on toxin type (Strotmeier et al. 2010, 2011; Rummel et al. 2003; Angstrom et al. 1994) on intestinal epithelial cells (Figure 1.11). One of the most exciting recent breakthroughs is the discovery that E-cadherin (E-cad), an epithelial cell surface adhesion molecule, is a host receptor for HA proteins. It was suggested that HA proteins directly bind to and

disrupt E-cad–mediated cell-to-cell adhesion, thus facilitating the absorption of BoNT

through the intestinal epithelium via a paracellular route (Ito et al. 2011; Matsumura et al.

2008; Sugawara et al. 2010). BoNTs thus stabilized as PTCs traverse the GI tract and

enter the bloodstream before they reach motor neurons.

   BoNT bind highly specific to non-myelinated areas of cholinergic neuromuscular

junctions (NMJs) (Dolly et al. 1984). Two synaptic vesicle proteins have been recently

identified as the neuronal receptors, namely, Synaptogamin (isoforms I and II) (Nishiki et

al., 1994, 1996a, 1996b) and Synaptic Vesicle 2(SV2) proteins (Dong et al., 2003, 2006,

2007, 2008). Therefore, BoNTs exploit this pathway for their internalization into neurons

by binding to the luminal domain of specific vesicle proteins (Montal, et al.,2010; Blum,

et al., 2012). The two domains LC and HC remain attached by the disulfide bond when

they are internalized into vesicles (Figure 1.8). Once internalized into the synaptic vesicle

lumen, acidification driven by the vATPase triggers the translocation of LC into the nerve

terminal cytoplasm that leads to the cleavage of SNARE proteins essential for membrane

fusion resulting in inhibition of neurotransmitter release at the NMJ (Montecucco et al.

1986, 1989; Montal 2010). Here, BoNTs act as sequence-specific endopeptidases that

cleave soluble N-ethylmaleimide sensitive factor attachment protein receptors

(SNAREs), blocking the release of acetylcholine at neuromuscular junctions (NMJs) and

thus paralyzing the affected muscles (Schiavo et al. 1992). The catalytic domain of

BoNTs is the LC which contains the active site of the neurotoxin and displays a very

specific metalloprotease activity (Schiavo et al. 2000). To date, the synaptic SNARE

(soluble NSF attachment protein receptor) proteins syntaxin-1-2-3, SNAP-25

(synaptosomal-associated protein of 25 kDa) and VAMP-1-2-3 (vesicle associated membrane protein, but also termed synaptobrevin-1, -2 and -3) are the only identified substrates of BoNT (Schiavo, et al., 2000). Cleavage of these synaptic SNARE proteins yields the persistent blockade of neurotransmitter release in intoxicated neurons (Schiavo et al. 1992; Blasi et al. 1993a, b; Schiavo et al. 1993a, b; Yamasaki et al. 1994). BoNT/A and BoNT/E cleave SNAP-25 while serotypes BoNT/B, BoNT/D, BoNT/F, and BoNT/G cleave synaptobrevin. BoNT/C is unique in that it is able to hydrolyze both syntaxin and SNAP-25 (Schiavo et al. 1995) (Figure 1.8). Once the toxin is endocytosed, the next step is to deliver LC across the intracellular membrane to its targets in the cytosol (Schiavo et al. 2000)

4. Mosquitocidal *Clostridium*: *Clostridium bifermentans* serovar *malaysia*

In 1990, WHO/TDR (WHO special program for research and training in tropical diseases) recommended and financed a nationwide screening program in Malaysia with the aim to identify new strains of insecticidal bacteria and other indigenous microbial control agents from soil and water samples. It led to the description of the first anaerobic isolate, a *Clostridium bifermentans* strain by Dr. L. H. Lim and S. Benjamin at the Institute for Medical Research, Kuala Lumpur, having a high mosquitocidal activity (Seleena et al., 1998). Further characterization and the existence of a specific H-antigen CH18 allowed it to be individualized as a serovar- *malaysia*. This strain was isolated from a mangrove swamp soil from Malaysia by W.C. Cheong and H.L. Lee and

identified *as C. bifermentans* serovar *malaysia* (de Barjac et al., 1990; Lee and Seleena, 1990a) at Institute de Pasteur.

Cbm toxicity against *A. stephensi* is about 10 times higher, but its toxicity against *Culex pipiens* or *Aedes aegypti* is 10 times lower, than that of Bti (Thiery *et al.,* 1992b). Larvicidal activity of Cbm was shown to be maximal during the sporulation 8 hours after beginning of the sporulation process (Figure 1.9), and in part was thought to be related to the presence of parasporal inclusion bodies (Charles *et al.,* 1990). After cell lysis, the larvicidal activity decreased by 10 times and this was due to the inactivation by extracellular proteases excreted in the supernatant. Different routes of administration like oral, percutaneous, subcutaneous, intraperitoneal or intravenous injections show the innocuity of the sporulated cells to mammals and goldfish (Thiery *et al.,* 1992). Other studies for Cbm toxicity against mammals and various other non-target organisms have also shown a lack of toxicity (Yiallouros *et al.,* 1994). These characteristics make Cbm a promising target for identification of a biological insecticide.

Unlike Bti, Cbm produces no parasporal inclusions that could be associated with toxicity (Charles *et al.,* 1990), and the toxic components have not been isolated. They were found to be unstable, being inactivated by physical filtration, sonication, or a cycle of freezing thawing (Nicolas *et al.,* 1990). Heating at 80 °C or by treatment with 50 mM NaOH destroyed the larvicidal activity whereas freeze-drying preserved it. Also no plasmid was found in this bacterium (Seleena *et al.,* 1994).

Previous attempts to characterize the toxic components led to the identification of four major proteins in crude cultures. Those putative toxins include a doublet of 66-68

kDa (*Cbm71* and *Cbm72* genes) and two other small proteins of 18 and 16 kDa (*Cbm17.1* and Cbm17.2) (Nicolas *et al.,.*, 1993a; Nicolas *et al.,* 1993b). The amino acid sequence of Cbm71 showed strong similarity to regions within the toxic domain of *B. thuringiensis* delta-endotoxins, i.e., the amino-terminal part (Lereclus *et al.,* 1993), especially within blocks I to IV. The sequence of block V and the carboxy terminus of delta-endotoxins are absent. These genes were renamed *Cry16A* and *Cry17A* (Barloy *et al.,* 1996) to indicate their relationship with the Cry-like toxin family. These Cry-like toxins are the first one found in an anaerobic organism. *Cbm17.1* and *Cbm17.2* genes showed low amino acid similarity (44%) to a hemolysin from *Aspergillus fumigatus* and retained their original nomenclature (Barloy *et al.,* 1998a).

These four proteins were immunologically unrelated to Bti or Ls toxins and were determined to belong to a novel class of insecticidal toxins (Nicolas *et al.,* 1990). In some mosquitocidal *Bacillus thuringiensis* strains, Cry toxins and hemolysins act synergistically, resulting in high levels of toxicity (Delécluse *et al.,* 2000). The presence of both Cry proteins and putative hemolysins in Cbm suggested a similar mode of action. Attempts were made to determine the role in toxicity of each of the proteins, by expression of the *Cry16A*, *Cry17A*, *Cbm 17.1*, and *Cbm 17.2* genes, alone or in combination, and under the control of their own promoters, in a crystal-minus Bt strain (Barloy *et al.,* 1996, 1998). Only the *Cry16A* gene was expressed and at a low level under these conditions. To overcome these problems, attempts were made to increase the expression of these four genes, either in *E. coli*, by use of the N-terminal His6x-tag system, or in Bt, in which the *Cry16A* and *Cry17A* genes were placed under the control of

23

the *Cry1C* gene promoter. Only *Cry16A* gene was successfully expressed in both systems but not secreted. Furthermore, recombinant strains of *E. coli* and Bt expressing Cry16A showed no toxicity to *Aedes aegypti* and *Anopheles stephensi*. No toxin production was detected in either Bt or *E. coli* transformed with Cry17A. In contrast, the Cry17A protein was detected by Western blot only when the *Cry16A* gene was co expressed in the same *Bt* strain. The recombinant *Bt* strain producing both Cry16A and Cry17A was totally inactive against mosquito larvae. The proteins produced by these cloned genes were not mosquitocidal (Juárez-Pérez and Delécluse, 2001), suggesting that the proteins responsible for toxicity to *Anopheles* mosquitoes remain unidentified.

5. Specific objectives of this study

Recognizing the potential use of Cbm strain for control of *Anopheles*, we started this project to identify genes and proteins in Cbm that have larvicidal activity against *Anopheles sp*. Our lab has demonstrated that the Cbm contains eight plasmids, six of which are not present in the Cb. As many toxin genes in many gram positive bacteria are encoded on plasmids including in Bti, we hypothesize that the toxin genes in Cbm are also encoded by these plasmids. Also in our lab, we have developed three mutant strains of Cbm by gamma irradiation that have lost their larvicidal activity. We have sequenced one of the mutant strains which show the absence of the largest plasmid of 109kb. This plasmid now designated pClosMP encodes for many known toxin genes.

Therefore, to understand Cbm toxicity in mosquitoes, further studies are needed to identify the genes involved in toxicity and should be examined in *Aedes*

aegypti, *Anopheles gambiae*, and *Culex quinquefasciatus*. This dissertation focuses on the cloning and expression of genes in two toxic operons, Cry toxin operon and Cmp toxin operon and investigate whether these proteins are individually or as a complex are required for mosquitocidal activity of Cbm in *Ae. aegypti*, An. *gambiae*, and *C. quinquefasciatus.*

Hypothesis: Plasmids confer mosquitocidal activity to Cbm.

Hypothesis 1: The pClosMP plasmid confers mosquitocidal activity to Cbm

A comparative genomic approach is taken to identify larvicidal toxin genes from Cbm. We sequenced the entire genomes of Cb, Cbm and Cbp. Since Cbm and Cbp are strains of Cb, the structure and organization of the genomes is very similar. This is helpful because clostridial genomes tend to vary a lot. It is also fortunate that we have Cb, which is non larvicidal. Our lab also developed three mutant strains of Cbm by gamma irradiation that have lost their larvicidal activity. Based on these results, I compared and analyzed the genomic sequences of Cbm, Cb and Cbp to search for the presence and absence of known and putative genes, as well as for the regions of synteny and divergence. I also wanted to get a complete genomic sequence of Cbm. Upon analysis of sequencing data, I found that the genome sizes of Cb and Cbm were little different. Size of Cb genome was about 3.6 MB whereas the size of Cbm genome was about 3.9 MB. This difference in size could be explained by the absence of 6-8 plasmids in Cb. Six plasmids present exclusively in Cbm were analyzed to identify candidate toxin genes for Cbm.

Aim 1.1: Analyze and compare the genomic sequences of Cb, Cbm and Cbp to identify the mosquitocidal genes in Cbm.

Aim 1.2: Complete the genomic and plasmid sequences of Cbm.

Hypothesis 2: The genes on the pClosMP plasmid are involved in Cbm mosquitocidal activity.

As mentioned above, six plasmids present exclusively in Cbm were analyzed to identify candidate toxin genes for Cbm. After analyzing BGI annotation, large plasmid of approx. 109 Kb showed the presence of the putative Cry16A, Cry17A, Cbm17.1 and Cbm17.2 genes in an operon which was earlier hypothesized or implicated in toxicity of Cbm (Barloy et al., 1996, 1998). Based on previous research and our genomic analysis, we cloned the complete native operon in Bt expression vector and tested the expression and toxicity on *Ae. aegypti*. Moreover, I investigated which key genes are involved in toxicity, by cloning and expressing individual gene and deletion operon constructs of the Cry toxin operon. Their mosquitocidal activity was checked in *Ae. aegypti*, *An. gambiae*, and *Culex quinquefasciatus*. On our investigation of the plasmid sequence, I had also found botulinum – like gene and several conserved associated genes in an operon flanked by transposons, in the plasmid. All the components of the human BoNT gene operon are present on the plasmid with promoter regions and Shine Delgarno sequences which should be expressed in this *Clostridium* strain. Hence, I investigated the mosquitocidal activity of these proteins expressed by themselves or in combination of 2 or all three

expressed together. Their mosquitocidal activity was also checked in *Ae. aegypti*, *An. gambiae* and *Culex quinquefasciatus.*

Aim 2.1: Test whether proteins in Cry toxin operon are individually or as a complex are required for mosquitocidal activity of Cbm

Aim 2.2: Test whether proteins in Cmp toxin operon are required for mosquitocidal activity of Cbm

Hypothesis 3: Cry17A plays an important role in toxicity of Cry Operon against *Aedes aegypti*.

The predicted structures of Cry16A and Cry17A have 34% in sequence homology to Cry8Ea1 and Cry Cry3Aa respectively and showed that they are composed of three distinct domains. The predicted structures of Cbm17.1 and Cbm17.2 have 33% in sequence homology to Cry34Ab1. One of the deletion mutants of Cry Operon, pCryO/Δ17 with a 636bp MscI - ApaI deletion in the Cry17A lost activity against *Aedes*. I investigated the role of this deleted domain in Cry Operon activity. I found that this deletion resulted in the deletion of domain I in Cry17A. As mentioned above, helices in Domain I are correlated to pore formation in some cases, and mutations/deletions in their sequence affected toxicity. Exposed loop regions in domain II have been shown to be important determinants of insect specificity (Bravo et al., 2005). Cry4Ba shows no toxicity against *Culex* sp while Cry4Aa is active against this mosquito species (Abdullah et al., 2003). It has been shown that by introducing domain II loop 3 Cry4Aa amino acid sequence into loop 3 of domain II of Cry4Ba resulted in a mutant toxin with toxicity

against *Culex* sp, retaining its insecticidal activity against *Aedes aegypti* (Abdullah et al., 2003). Therefore, I investigated the possibility that exchanging pore forming helices from one Cry toxin to another pore forming Cry toxin that shows toxicity to mosquitoes (like Cry4Ba) might also transfer the toxicity determinants as well. For this experiment, I chose to swap the deleted portion of Cry17A with that of mosquitocidal Cry4Ba toxin as it is also a pore forming toxin and my lab already had a stable subclone in *E. coli* and Bt 4Q7 cells. The objective of the Domain I exchange experiment was to re- introduce *Aedes* toxicity to the CryΔ17A mutant of the Cry Operon. I had shown that Cbm Cry Operon is very toxic to *Aedes* but not toxic at all to *Anopheles*, and *Culex larvae* (Qureshi et al., 2014). Cry4Ba has been reported to be very toxic to *Aedes* and *Anopheles* larvae but not toxic at all to *Culex* larvae (Angsuthanasombat et al., 1992). I also investigated the possibility that exchanging loop residues from Cry17A toxin to Cry4Ba toxin might also transfer the toxicity determinants as well. The objective of the loop exchange experiment was to introduce *Anopheles* and *Culex* toxicity into Cry Operon by matching as closely as possible to the apical loop sequences and the length of the loops in Cry4Aa and Cry4Ba by mutagenesis.

Aim 3.1: Domain I swap of Cry17A with Cry4Ba as an approach to reclaiming toxicity against *Aedes aegypti*.

Aim 3.2: Domain II loop exchanges as a rational approach to modifying specificity

References

Abdullah, M. A., Valaitis, A. P. and Dean, D. H. (2006). Identification of a *Bacillus thuringiensis* Cry11Ba toxin-binding aminopeptidase from the mosquito, *Anopheles quadrimaculatus*. *BMC Biochem* **7**: 16.

Abdullah, M. A., 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. *AEM*, **69**(9): 5343–5353

Angsuthanasombat, C., Crickmore, N. and Ellar, D. J. (1991). Cytotoxicity of a cloned *Bacillus thuringiensis* subsp, *israelensis* CryIVB toxin to an *Aedes aegypti* cell line. *FEMS Microbiol Lett.***83**: 273-276.

Allured, V. S., R. J. Collier, S. F. Carroll, and D. B. McKay. (1986). Structure of exotoxin A of *Pseudomonas aeruginosa* at 3.0-Ångstrom resolution. *Proc. Natl. Acad. Sci. USA* **83**:1320-1324.

Ahnert-Hilger, G. and Bigalke, H. (1995). Molecular aspects of tetanus and botulinum neurotoxin poisoning.*Prog Neurobiol*. **46**(1): 83-96.

Amatsu, S., Sugawara, Y., Matsumura, T., Kitadokoro, K. and Fujinaga, Y. (2013). Crystal structure of *Clostridium botulinum* whole hemagglutinin reveals a huge triskelion-shaped molecular complex. *J Biol Chem*. **288**:35617–35625

Angus, T. A. (1956). Association of toxicity with protein-crystalline inclusions of Bacillus sotto Ishiwata. *Can J Microbiol* **2**: 111-121.

Angstrom, J., Teneberg, S., Karlsson, K. A. (1994). Delineation and comparison of ganglioside-binding epitopes for the toxins of *Vibrio cholerae*, *Escherichia coli*, and *Clostridium tetani*: evidence for overlapping epitopes. *Proc Natl Acad Sci USA* **91**(25):11859–11863

Arndt, J.W., Gu, J., Jaroszewski, L., Schwarzenbacher, R.,Hanson, M.A., Lebeda, F.J. and Stevens, R.C. (2005). The structure of the neurotoxin-associated protein HA33/A from *Clostridium botulinum* suggests a reoccurring beta-trefoil fold in the progenitor toxin complex. *J Mol Bio.* **346**:1083–1093

Barash, J.R. and Arnon, S.S. (2004). Dual toxin-producing strain of *Clostridium botulinum* type Bf isolated from a California patient with infant botulism. *J. Clin. Microbiol*. **42**, 1713–1715.

Barloy, F., Delécluse, A., Nicolas, L. and Lecadet, M. M. (1996). Cloning and expression of the first anaerobic toxin gene from *Clostridium bifermentans* serovar *malaysia*,

encoding a new mosquitocidal protein with homologies to *Bacillus thuringiensis* delta - endotoxins. *J. Bacteriol.* **178**: 3099–3105.

Barloy, F., Lecadet, M. M. and Delécluse, A. (1998). Cloning and sequencing of three new putative toxin genes from *Clostridium bifermentans* CH18. *Gene* **211:** 293–299.

Barrett, A. D. and Higgs, S. (2007). Yellow fever: a disease that has yet to be conquered**.** *Annu Rev Entomol.* **52:** 209-29.

Baxter, S. W, Badenes-Pérez, F. R., Morrison, A., Vogel, H., Crickmore, N., Kain, W., Wang, P., Heckel, D. G. and Jiggins, C. D. (2011). Parallel evolution of *Bacillus thuringiensis* toxin resistance in lepidoptera. *Genet*. **189**:675–679.

Bayyareddy, K., Andacht, T. M., Abdullah, M. A. and Adang, M. J. (2009). Proteomic identification of *Bacillus thuringiensis* subsp. *israelensis* toxin Cry4Ba binding proteins in midgut membranes from *Aedes* (Stegomyia) *aegypti* Linnaeus (Diptera, Culicidae) larvae. *Insect Biochem Mol Biol* **39**: 279-286.

Blasi, J., Chapman, E.R., Link, E., Binz, T., Yamasaki, S., De Camilli, P., Südhof, T. C., Niemann, H. and Jahn, R. (1993a). Botulinum neurotoxin A selectively cleaves the synaptic protein SNAP-25. *Nature*. **365**:160–163

Blasi, J., Chapman, E. R., Yamasaki, S., Binz, T., Niemann, H. and Jahn, R. (1993b). Botulinum neurotoxin C1 blocks neurotransmitter release by means of cleaving HPC-1/syntaxin. *EMBO J* **12**:4821–4828.

Blum, F. C., Chen, C., Kroken, A. R. and Barbieri, J. T. (2012). Tetanus toxin and botulinum toxin a utilize unique mechanisms to enter neurons of the central nervous system. *Infect Immun* **80**: 1662–1669.

Benefield, A., Dessain, S. K., Shine, N., Ohi, M.D. and Lacy, B. (2013). Molecular assembly of botulinum neurotoxin progenitor complexes. *Proc Natl Acad Sci U S A* **110**:5630–5635.

Berry, C., O'Neil, S., Ben-Dov, E., Jones, A. F., Murphy, L., Quail, M. A., Holden, M. T., Harris, D., Zaritsky, A. and Parkhill, J. (2002). Complete sequence and organization of pBtoxis, the toxin-coding plasmid of *Bacillus thuringiensis* subsp. *israelensis*. *Appl Environ Microbiol* **68**: 5082-5095.

Binz, T., Kurazono, H., Wille, M., Frevert, J., Wernars, K., and Niemann, H. (1990a). The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins. *J Biol Chem.* **265**(16), 9153-9158.

Binz, T., Kurazono, H., Popoff, M.R., Eklund, M.W., Sakaguchi, G., Kozaki, S., Krieglstein, K., Henschen, A., Gill, D.M. and Niemann, H. (1990b). Nucleotide sequence of the gene encoding *Clostridium botulinum* neurotoxin type D. *Nucleic Acids Res.* **18**(18), 5556.

Binz, T., Bade, S., Rummel, A., Kollewe, A. and Alves, J. (2002). Arg362 and Tyr365 of the botulinum neurotoxin type A light chain are involved in transition state stabilization. *Biochemistry* . **41**, 1717-1723.

Binz, T. and Rummel, A. (2009) .Cell entry strategy of clostridial neurotoxins. *J Neurochem.* **109**, 1584-1595.

Boonserm, P., Davis, P., Ellar, D. J. and Li, J. (2005). Crystal structure of the mosquito-larvicidal toxin Cry4Ba and its biological implications. *J Mol Biol* **348**: 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. *J Bacteriol* **188**: 3391-3401.

Bravo, A., Gill S. S. and Soberon M. (2007). Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon.* **49** (4): 423-435.

Briggs, D.C., Naylor, C.E., Smedley, J. G., Lukoyanova, N., Robertson, S., Moss, D.S., McClane, B.A. and Basak, A.K. (2011). Structure of the Food-Poisoning *Clostridium perfringens* Enterotoxin Reveals Similarity to the Aerolysin- Like Pore-Forming Toxins. *J Mol Biol.* **413** (1):138-49.

Butko, P. (2003). Cytolytic Toxin Cyt1A and its Mechanism of Membrane Damage: Date and Hypotheses. *Appl Environ Microbiol.* **69**:2415-2422.

Campbell, K., Collins, M. D. and East, A. K. (1993). Nucleotide sequence of the gene coding for *Clostridium botulinum* (*Clostridium argentinense*) type G neurotoxin: genealogical comparison with other clostridial neurotoxins. *Biochim Biophys Acta* .**1216**(3): 487-91.

Chai, Q., Arndt, J.W., Dong, M., Tepp, W.H., Johnson, E.A., Chapman, E.R. and Stevens, R.C. (2006). Structural basis of cell surface receptor recognition by botulinum neurotoxin B. *Nature.* **444**, 1096-100.

Charles, J. -F., L. Nicolas, M. Sebald, and H. de Barjac. (1990). *Clostridium bifermentans* serovar *malaysia*: sporulation, biogenesis of inclusion bodies and larvicidal effect on mosquito. *Res. Microbiol.* **141:** 721–733.

Charles, J.-F., and Nielsen-LeRoux, C. (2000). Mosquitocidal bacterial toxins: Diversity, mode of action and resistance phenomena. *Mem Institut Oswaldo Cruz.* **95**: 201-206.

Chen, J., Aimanova, K. G., Fernandez, L. E., Bravo, A., Soberon, M. and Gill, S. S. (2009). *Aedes aegypti* cadherin serves as a putative receptor of the Cry11Aa toxin from *Bacillus thuringiensis* subsp. *israelensis*. *Biochem J* **424**: 191-200.

Chen, J., Aimanova, K. G., Pan, S. and Gill, S. S. (2009). Identification and characterization of *Aedes aegypti* aminopeptidase N as a putative receptor of *Bacillus thuringiensis* Cry11A toxin. *Insect Biochem Mol Biol* **39**: 688-696.

Choe, S., M. J. Bennett, G. Fujii, P. M. G. Curmi, K. A. Kantardjieff, R. J. Collier, and D. Eisenberg. (1992). The crystal structure of diphtheria toxin. *Nature*. **357**:216-222.

Cohen, S., Albeck, S., Ben-Dov, E., Cahan, R., Firer, M., Zaritsky, A. and Dym, O. (2011). Cyt1Aa toxin: crystal structure reveals implications for its membrane-perforating function. *J Mol Biol.* **413**: 804-814.

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. *Microbiol Mol Biol Rev* **62**: 807-813.

Crickmore, N., Zeigler, D.R., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J., Bravo, A. and Dean, D.H. (2011). *Bacillus thuringiensis* toxin nomenclature. http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/index.html

Collins, M.D. and East, A.K. (1998). Phylogeny and taxonomy of the foodborne pathogen *Clostridium botulinum* and its neurotoxins. *J Appl. Microbiol.* **84**, 5-17.

Couesnon, A., Raffestin, S. and Popoff, M. R. (2006). Expression of botulinum neurotoxins A and E, and associated non-toxin genes, during the transition phase and stability at high temperature: analysis by quantitative reverse transcription-PCR. *Microbiology.* **152**, 759-70.

Cox, F.E.G. (2002). History of human parasitology. *Clin Microbiol Rev* **15:** 595-612.

Davies, A.H., Roberts, A.K., Shone, C.C. and Acharya, K.R. (2011). Super toxins from a super bug: structure and function of *Clostridium difficile* toxins. *Biochem J .***436**(3), 517-26.

de Barjac, H., Sebald, M., Charles, J.-F., Cheong, W. H. and Lee, H. H. (1990). *Clostridium bifermentans* serovar malaysia, une nouvelle bacte ´rie anae´robie pathoge `ne des larves de moustiques et de simulies. *C.R. Acad. Sci. Se´r. III* **310:** 383–387.

de Maagd, R. A., Bravo, A. and Crickmore, N. (2001). How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. *Trends Genet* **17**: 193-199.

Dechklar, M., Tiewsiri, K., Angsuthanasombat, C. and Pootanakit, K. (2011). Functional expression in insect cells of glycosylphosphatidylinositol-linked alkaline phosphatase from *Aedes aegypti* larval midgut: a *Bacillus thuringiensis* Cry4Ba toxin receptor. *Insect Biochem Mol Biol.* **41**: 159-166.

Delécluse, A., Juárez-Pérez, V. and Berry, C. (2000). Vector-active toxins: structure and diversity, p. 101-125. *In* Entomopathogenic Bacteria: from Laboratory to Field Application. Springer ebook.

Dhurua, S. and Gujar, G.T. (2011). Field-evolved resistance to Bt toxin Cry1Ac in the pink bollworm, *Pectinophora gossypiella* (Saunders) (Lepidoptera: Gelechiidae), from India. *Pest Manag Sci* **67**: 898–903.

Dolly, J. O., Black, J., Williams, R. S. and Melling, J. (1984). Acceptors for botulinum neurotoxin reside on motor nerve terminals and mediate its internalization. *Nature*. **307**(5950):457–60

Dong, M., Richards, D. A., Goodnough, M. C., Tepp, W. H., Johnson, E. A. and Chapman, E. R. (2003). Synaptotagmins I and II mediate entry of botulinum neurotoxin B into cells. *J Cell Biol*. **162**(7):1293–1303

Dong, M., Yeh, F., Tepp, W.H., Dean, C., Johnson, E. A., Janz, R. and Chapman, E. R. (2006). SV2 is the proteinreceptor for botulinum neurotoxin A. *Science*. **312**(5773):592–596

Dong, M., Tepp, W. H., Liu, H., Johnson, E. A. and Chapman, E.R. (2007). Mechanism of botulinum neurotoxin B and G entry into hippocampal neurons. *J Cell Biol* **179**(7):1511–1522

Dong, M., Liu, H., Tepp, W. H., Johnson, E. A., Janz, R. and Chapman, E. R. (2008). Glycosylated SV2A and SV2B mediate the entry of botulinum neurotoxin E into neurons. *Mol Biol Cell.* **19**(12):5226–5237

Ellis R. T., Stockhoff B. A., Stamp L., Schnepf H. E., Schwab G. E., Knuth M., Russell J., Cardineau G. A. and Narva K. E. (2002). Novel *Bacillus thuringiensis* binary insecticidal crystal proteins active on western corn rootworm, *Diabrotica virgifera virgifera* LeConte. *Appl. Environ. Microbiol.* **68**(3): 1137-1145.

Estruch, J.J., Warren, G.W., Mullins, M.A., Nye, G.J., Craig, J.A. and Koziel, M.G. (1996). Vip3A, a novel *Bacillus thuringiensis* vegetative insecticidal protein with a wide spectrum of activities against lepidopteran insects. *PNAS*. **93**: 5389-5394.

Feany, M. B., Lee, S., Edwards, R.H. and Buckley, K. M. (1992) The synaptic vesicle protein SV2 is a novel type of transmembrane transporter. *Cell* **70**(5):861–867

Ffrench-Constant, R.H. and Bowen, D.J. (2000). Novel insecticidal toxins from nematode-symbiotic bacteria. *Cell Mol Life Sci* **57**: 828–833.

Fernandez, L. E., Aimanova, K. G., Gill, S. S., Bravo, A. and Soberon, M. (2006). A GPI-anchored alkaline phosphatase is a functional midgut receptor of Cry11Aa toxin in *Aedes aegypti* larvae. *Biochemical J* **394**: 77-84.

Fujinaga, Y., Inoue, K., Watanabe, S., Yokota, K., Hirai, Y., Nagamachi, E. and Oguma K. (1997). The haemagglutinin of *Clostridium botulinum* type C progenitor toxin plays an essential role in binding of toxin to the epithelial cells of guinea pig small intestine, leading to the efficient absorption of the toxin. *Microbiology* .**143**, 3841–3847

Fujinaga, Y., Sugawara, Y. and Matsumura, T. (2013). Uptake of Botulinum Neurotoxin in the Intestine. *Curr Topic Microbiol. Immunol.* **364**:45-59

Gahan, L. J., Pauchet, Y., Vogel, H. and Heckel, D. G. (2010). An ABC transporter mutation is correlated with insect resistance to *Bacillus thuringiensis* Cry1Ac toxin. *PLoS Genet.* **6**:e1001248.

Galitsky, N., Cody,v., Wojtczak, A. , Ghosh, D., Luft, J. R., Pangborn, W. and English., L. (2001). Structure of the insecticidal bacterial delto-endotoxin Cry3Bb1 of *Bacillus thuringiensis*. *Acta Crystallogr D Biol Crystallogr* **57**:1101-9.

Gardiner, D. L., McCarthy, J. S. and Trenholme, K. R. (2005). Malaria in the post-genomics era: light at the end of the tunnel or just another train? *Postgrad Med J.* **81**(958): 505-9.

Gassmann, A.J., Petzold-Maxwell, J.L., Keweshan, R.S. and Dunbar, M.W. (2011). Field-evolved resistance to Bt maize by western corn rootworm. *PLoS ONE* **6**: e22629.

Girardin, H., Albagnac, C., Dargaignaratz, C., Nguyen-The, C. and Carlin F. (2002) Antimicrobial activity of foodborne *Paenibacillus* and *Bacillus* spp. against *Clostridium botulinum*. *J Food Prot.* **65**:806-813.

Goldberg, L. J. and Margalit, J. (1977). A bacterial spore demonstrating rapid larvicidal activity against Anopheles sergentii, Uranotaenia unguiculata, Culex univitattus, Aedes aegypti and Culex pipiens. Mosq News 37: 355-358.

Gomez I., Pardo-Lopez L., Munoz-garay C., Fernandez L. E., Pérez C., Sanchez J., Soberon M. and Bravo, A. (2007). Role of receptor interaction in the mode of action of

Feany, M. B., Lee, S., Edwards, R.H. and Buckley, K. M. (1992) The synaptic vesicle protein SV2 is a novel type of transmembrane transporter. *Cell* **70**(5):861–867

Ffrench-Constant, R.H. and Bowen, D.J. (2000). Novel insecticidal toxins from nematode-symbiotic bacteria. *Cell Mol Life Sci* **57**: 828–833.

Fernandez, L. E., Aimanova, K. G., Gill, S. S., Bravo, A. and Soberon, M. (2006). A GPI-anchored alkaline phosphatase is a functional midgut receptor of Cry11Aa toxin in *Aedes aegypti* larvae. *Biochemical J* **394**: 77-84.

Fujinaga, Y., Inoue, K., Watanabe, S., Yokota, K., Hirai, Y., Nagamachi, E. and Oguma K. (1997). The haemagglutinin of *Clostridium botulinum* type C progenitor toxin plays an essential role in binding of toxin to the epithelial cells of guinea pig small intestine, leading to the efficient absorption of the toxin. *Microbiology* .**143**, 3841–3847

Fujinaga, Y., Sugawara, Y. and Matsumura, T. (2013). Uptake of Botulinum Neurotoxin in the Intestine. *Curr Topic Microbiol. Immunol.* **364**:45-59

Gahan, L. J., Pauchet, Y., Vogel, H. and Heckel, D. G. (2010). An ABC transporter mutation is correlated with insect resistance to *Bacillus thuringiensis* Cry1Ac toxin. *PLoS Genet.* **6**:e1001248.

Galitsky, N., Cody,v., Wojtczak, A. , Ghosh, D., Luft, J. R., Pangborn, W. and English., L. (2001). Structure of the insecticidal bacterial delto-endotoxin Cry3Bb1 of *Bacillus thuringiensis*. *Acta Crystallogr D Biol Crystallogr* **57**:1101-9.

Gardiner, D. L., McCarthy, J. S. and Trenholme, K. R. (2005). Malaria in the post-genomics era: light at the end of the tunnel or just another train? *Postgrad Med J.* **81**(958): 505-9.

Gassmann, A.J., Petzold-Maxwell, J.L., Keweshan, R.S. and Dunbar, M.W. (2011). Field-evolved resistance to Bt maize by western corn rootworm. *PLoS ONE* **6**: e22629.

Girardin, H., Albagnac, C., Dargaignaratz, C., Nguyen-The, C. and Carlin F. (2002) Antimicrobial activity of foodborne *Paenibacillus* and *Bacillus* spp. against *Clostridium botulinum*. *J Food Prot.* **65**:806-813.

Goldberg, L. J. and Margalit, J. (1977). A bacterial spore demonstrating rapid larvicidal activity against Anopheles sergentii, Uranotaenia unguiculata, Culex univitattus, Aedes aegypti and Culex pipiens. Mosq News 37: 355-358.

Gomez I., Pardo-Lopez L., Munoz-garay C., Fernandez L. E., Pérez 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): 169-173.

Gonzalez, J. M., Jr., Brown, B. J. and Carlton, B. C. (1982). Transfer of *Bacillus thuringiensis* plasmids coding for delta-endotoxin among strains of *B. thuringiensis* and *B. cereus*. *Proc Natl Acad Sci U S A* **79**: 6951-6955.

Gu, S., Rumpel, S., Zhou, J., Strotmeier, J., Bigalke, H., Perry, K., Shoemaker, C.B., Rummel, A. and Jin, R. (2012). Botulinum neurotoxin is shielded by NTNHA in an interlocked complex. *Science* . **335**(6071), 977–981.

Gubler, D. J. and Casta-Valez, A. (1991). A program for prevention and control of epidemic dengue and dengue hemorrhagic fever in Puerto Rico and the U.S. Virgin Islands. *Bull Pan Am Health Organ.* **25(3):** 237-47.

Gubler, D. J. (1997). Human behaviour and cultural context in disease control. *Trop Med Int Health.* **2(11):** A1-2.

Gubler, D. J. (1998). Resurgent vector-borne diseases as a global health problem. *Emerg Infect Dis* **4**: 442-450.

Gubler, D. J. (2004). The changing epidemiology of yellow fever and dengue, 1900 to 2003: full circle? *Comp Immunol Microbiol Infect Dis* **27:** 319-30.

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*. *J. Struct. Biol.* **168**, 259–266.

Hasegawa, K., Watanabe, T., Suzuki, T., Yamano, A., Oikawa, T., Sato, Y., Kouguchi, H., Yoneyama, T., Niwa, K., Ikeda, T. and Ohyama, T. (2007). A novel subunit structure of *Clostridium botulinum* serotype D toxin complex with three extended arms. *J Biol Chem.* **282**:24777–24783.

Hatheway, C.L. (1990) Toxigenic clostridia. *Clin. Microbiol. Rev.***3**, 66-98.

Heckel D. G. (2012). Learning the ABCs of Bt: ABC transporters and insect resistance to *Bacillus thuringiensis* provide clues to a crucial step in toxin mode of action. *Pest. Biochem. Physiol.* **104** (2): 103-110.

Helgason, E., Caugant, D. A., Lecadet, M.-M., Chen, Y., Mahillon, J., Lövgren,A., Hegna, I., Kvaløy, K. and Kolstø, A.-B. (1998). Genetic diversity of *Bacillus cereus / Bacillus thuringiensis* isolates from natural sources. *Curr. Microbiol.* **37**:80-87.

Hill, K.K., Smith, T.J., Helma, C.H., Ticknor, L.O., Foley, B.T., Svensson, R.T., Brown, J.L., Johnson, E.A., Smith, L.A., Okinaka, R.T., Jackson, P.J. and Marks, J.D. (2007). Genetic diversity among Botulinum Neurotoxin-producing clostridial strains. *J Bacteriol.* **189**, 818-832.

Hofte, H. and Whiteley, H. R. (1989). Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiol Rev* **53**: 242-255.

Inoue, K., Fujinaga, Y., Honke, K., Arimitsu, H., Mahmut, N., Sakaguchi, Y., Ohyama, T., Watanabe, T., Inoue, K. and Oguma, K. (2001). *Clostridium botulinum* type A haemagglutinin-positive progenitor toxin (HA(+)-PTX) binds to oligosaccharides containing Gal beta1-4GlcNAc through one subcomponent of haemagglutinin (HA1). *Microbiology.* **147**:811–819.

Inoue, K., Sobhany, M., Transue, T.R,. Oguma, K., Pedersen, L.C. and Negishi, M. (2003). Structural analysis by X-ray crystallography and calorimetry of a haemagglutinin component (HA1) of the progenitor toxin from *Clostridium botulinum. Microbiology* **149**:3361–3370.

Ito, H., Sagane, Y., Miyata, K., Inui, K., Matsuo, T., Horiuchi, R., Ikeda, T., Suzuki, T., Hasegawa, K., Kouguchi, H., Oguma, K., Niwa, K., Ohyama, T. and Watanabe, T. (2011). HA-33 facilitates transport of the serotype D botulinum toxin across a rat intestinal epithelial cell monolayer. *FEMS Immunol Med Microbiol.* **61**:323–331.

Juárez-Pérez, V. and Delécluse, A. (2001). The Cry toxins and the putative hemolysins of *Clostridium bifermentans* ser. *malaysia* are not involved in mosquitocidal activity. *J Invertebr Pathol.* **78(1):** 57-8.

Jiménez, A. I., Reyes, E. Z., Cancino-Rodezno, A., Bedoya-Pérez, L. P., Caballero-Flores, G. G., Muriel-Millan, L. F., Likitvivatanavong, S., Gill, S. S., Bravo, A. and Soberón, M. (2012). *Aedes aegypti* alkaline phosphatase ALP1 is a functional receptor of *Bacillus thuringiensis* Cry4Ba and Cry11Aa toxins. *Insect Biochem Mol Biol* **42**: 683-689.

Katayama, S., Dupuy, B., Garnier, T. and Cole, S.T. (1995) Rapid expansion of the physical and genetic map of the chromosome of *Clostridium perfringens* CPN50. *J. Bacteriol.* **177**, 5680- 5685.

Kubota, T., Yonekura, N., Hariya, Y., Isogai, E., Isogai, H., Amano, K. and Fujii, N. (1998). Gene arrangement in the upstream region of *Clostridium botulinum* type E and *Clostridium butyricum* BL6340 progenitor toxin genes is different from that of other types. *FEMS Microbiol. Lett.* **158**, 215-221.

Kumaran, D., Eswaramoorthy, S., Furey, W., Navaza, J., Sax, M. and Swaminathan, S. (2009). Domain organization in *Clostridium botulinum* neurotoxin type E is unique: its implication in faster translocation. *J. Mol. Biol*. **386**, 233-245.

Kurazono, H., Mochida, S., Binz, T., Eisel, U., Quanz, M., Grebenstein, O., Wernars, K., Poulain, B., Tauc, L. and Niemann, H. (1992). Minimal essential domains specifying toxicity of the light chains of tetanus toxin and botulinum neurotoxin type A. *J Biol Chem*. **267**:14721–14729.

Lacy, D.B., Tepp, W., Cohen, A.C., DasGupta, B.R. and Stevens, R.C. (1998). Crystal structure of botulinum neurotoxin type A and implications for toxicity. *Nat. Struct. Biol*. **5**, 898-902.

Lacy, D.B. and Stevens, R.C. (1999). Sequence homology and structural analysis of clostridial neurotoxins. *J. Mol. Biol*. **291**, 1091-1104.

Lacey, L. A. (2007). *Bacillus thuringiensis* serovariety *israelensis* and *Bacillus sphaericus* for mosquito control. *J Am Mosq Control Assoc* **23**: 133-163.

Lanciotti, R. S., Roehrig, J. T., Deubel, V., Smith, J., Parker, M., Steele, K., Crise, B., Volpe, K. E., Crabtree, M. B., Scherret, J. H., Hall, R. A., MacKenzie, J. S., Cropp, C. B., Panigrahy, B., Ostlund, E., Schmitt, B., Malkinson, M., Banet, C., Weissman, J., Komar, N., Savage, H. M., Stone, W., McNamara, T. and Gubler, D. J. (1999). Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science*. **286(5448):** 2333-7.

Lee, H. P. and Seleena, P.(1990). Isolation of indigenous larvicidal microbial control agents of mosquitoes: The Malaysian experience. S*outheast Asian J. Trop. Med. Public Health*. **21**: 281–287.

Lee, S.B., Aimanova, K. and Gill, S.S. (2014). Alkaline phosphatases and aminopeptidases are altered in a Cry11Aa resistant strain of *Aedes aegypti*. *Insect Biochem Mol Biol*. **54**:112-21.

Lee, K., Gu, S., Jin, L., Le, T. T. T. N., Cheng, L. W., Strotmeier, J., Kruel, A. M., Yao, G., Perry, K., Rummel, A. and Jin, R. (2013). Structure of a bimodular botulinum neurotoxin complex provides insights into its oral toxicity. *PLOS Pathog* **9**:e1003690

Lereclus, D., A. Dele´cluse, and M.-M. Lecadet. (1993). Diversity of *Bacillus thuringiensis* toxins and genes. *In Bacillus thuringiensis* and environmental biopesticide: theory and practice. John Wiley & Sons, Ltd., London. p. 36–69

Li, J., Koni, P. A. and Ellar, D. J. (1996). Structure of the mosquitocidal delta-endotoxin CytB from *Bacillus thuringiensis* sp. *kyushuensis* and implications for membrane pore formation. *J Mol Biol* **257**: 129-152.

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**: 815-821.

Likitvivatanavong S. , Chen J., Bravo, M. Soberon, Gill S. S. (2011). Cadherin, alkaline phosphatase, and N-aminopeptidase as receptors of Cry11Ba toxin from *Bacillus thuringiensis* subsp.. *jegathesan* in *Aedes aegypti*. *Appl. Environ. Microbiol*. **77** (1): 24-31.

Lövenklev, M. (2003). A molecular approach for investigation of the prevalence and neurotoxin formation of *Clostridium botulinum* in food safety assessment. *Applied Microbiology*. Lund, Lund Institute of Technology.

Luthy, P. and Ebersold, H. R. (1981). *Bacillus thuringiensis* delta-endotoxin: Histopathology and molecular mode of action. In: Pathegenesis of Invertebrate Microbiol Disease. Davison, E. W. (Ed). pp235-267. Allenheld, Osmun & Co. Publishers, Inc., Totowa. N.J.

Ramírez-Lepe, M. and Ramírez-Suero, M. (2012). Biological Control of Mosquito Larvae by Bacillus thuringiensis subsp. israelensis, Insecticides - Pest Engineering, Dr. Farzana Perveen (Ed.)

McMahon HT, Boucrot E. (2011). Molecular mechanism and physiological functions of clathrinmediated endocytosis. Nat Rev Mol Cell Biol 12:517–533.

Marvaud, J. C., Gibert, M., Inoue, K., Fujinaga, Y., Oguma, K. and Popoff, M.R.(1998). botR/A is a positive regulator of botulinum neurotoxin and associated non-toxin protein genes in *Clostridium botulinum* A. *Mol Microbiol* . **29**(4), 1009-18.

Matsumura, T., Jin, Y., Kabumoto, Y., Takegahara, Y., Oguma, K., Lencer, W.I. and Fujinaga, Y. (2008). The HA proteins of botulinum toxin disrupt intestinal epithelial intercellular junctions to increase toxin absorption. *Cell Microbiol* .**10**(2), 355–364.

McNall, R. J. and Adang, M. J. (2003). Identification of novel *Bacillus thuringiensis* Cry1Ac binding proteins in *Manduca sexta* midgut through proteomic analysis. *Insect Biochem Mol Biol* **33**: 999-1010.

Melville, S. (2007). Clostridia: diarrheal disease, tissue infection, botulism, and tetanus. In *Schaechter's mechanisms of microbial disease*, Fourth edition. Edited by Engleberg NC, Dermody TS, DiRita V, Lippincott Williams & Wilikins, Philadelphia, Pennsylvania, 230-235.

Mittal, P. K. (2003). Biolarvicides in vector control: challenges and prospects. *J Vector Borne Dis* **40**: 20-32.

Montal, M. (2010). Botulinum neurotoxin: a marvel of protein design. *Annu Rev Biochem*. **79**:591–617

Montecucco, C. (1986). How do tetanus and botulinum toxins bind to neuronal membranes? *Trends Biochem Sci*. **11**:315–317

Montecucco, C., Schiavo, G. and Dasgupta, B. R. (1989). Effect of pH on the interaction of botulinum neurotoxins A, B and E with liposomes. *Biochem J* . **259**:47–53.

Montecucco, C., Rossetto, O. and Schiavo, G.(2004). Presynaptic receptor arrays for clostridial neurotoxins. *Trends Microbiol*. **12**(10):442-6.

Moriishi, K., Koura, M., Abe, N., Fujii, N., Fujinaga, Y., Inoue, K. and Oguma, K. (1996). Mosaic structures of neurotoxins produced from *Clostridium botulinum* types C and D organisms. *Biochim. Biophys. Acta* .**1307**, 123-126.

Morse R. J., Yamamoto T. and Stroud R. M. (2001). Structure of Cry2Aa suggests an unexpected receiver binding epitope. *Structure*. **9** (5): 409-417.

Nash, D., Mostashari, F., Fine, A., Miller, J., O'Leary, D., Murray, K., Huang, A., Rosenberg, A., Greenberg, A., Sherman, M., Wong, S. and Layton, M. (1999). The outbreak of West Nile virus infection in the New York City area in 1999. *N Engl J Med*. **344(24):** 1807-14.

Nakamura, T., Kotani, M., Tonozuka, T., Ide, A., Oguma, K. and Nishikawa, A. (2009). Crystal structure of the HA3 subcomponent of *Clostridium botulinum* type C progenitor toxin. *J Mol Biol*. **385**:1193–1206

Nakamura, T., Tonozuka, T., Ito, S., Takeda, Y., Sato, R., Matsuo, I., Ito, Y., Oguma, K. and Nishikawa, A. (2011). Molecular diversity of the two sugar-binding sites of the beta-trefoil lectin HA33/C (HA1) from *Clostridium botulinum* type C neurotoxin. *Arch Biochem Biophys*. **512**:69–77

Nicolas, L., Charles, J.-F., and de BarjaC, H. (1993). *Clostridium bifermentans* serovar *malaysia*: characterization of putative mosquito larvicidal proteins. *FEMS Microbiol. Lett.* **113:** 23–28.

Nicolas, L., Hamon, S., Frachon, E., Sebald, M. and de BarjaC, H. (1990). Partial inactivation of the mosquitocidal activity of *Clostridium bifermentans* serovar *malaysia* by extracellular proteinases. *Appl. Microbiol. Biotechnol.* **34:** 36–41.

Nielsen-Leroux, C., Charles, J.F., Thiery, I., and Georghiou, G.P. (1995). Resistance in a laboratory population of *Culex quinquefasciatus* (Diptera: Culicidae) to *Bacillus sphaericus* binary toxin is due to a change in the receptor on midgut brush-border membranes. *Eur J Biochem.* **228**: 206-210.

Nishiki, T., Kamata, Y., Nemoto, Y., Omori, A., Ito, T., Takahashi, M. Aand Kozaki, S. (1994). Identification of protein receptor for *Clostridium botulinum* type B neurotoxin in rat brain synaptosomes. *J Biol Chem* **269**(14):10498–10503.

Nishiki, T., Tokuyama, Y., Kamata, Y., Nemoto, Y., Yoshida, A., Sato, K., Sekiguchi, M., Takahashi, M. and Kozaki, S. (1996a). The high-affinity binding of *Clostridium botulinum* type B neurotoxin to synaptotagmin II associated with gangliosides GT1b/GD1a. *FEBS Lett.* **378**(3):253–257.

Nishiki, T., Tokuyama, Y., Kamata, Y., Nemoto, Y., Yoshida, A., Sekiguchi, M., Takahashi, M. and Kozaki, S. (1996b). Binding of botulinum type B neurotoxin to Chinese hamster ovary cells transfected with rat synaptotagmin II cDNA. *Neurosci Lett* **208**(2):105–108.

Nishikawa, A., Uotsu, N., Arimitsu, H., Lee, J-C., Miura, Y., Fujinaga, Y., Nakada, H., Watanabe, T., Ohyama, T., Sakano, Y. and Oguma K. (2004). The receptor and transporter for internalization of *Clostridium botulinum* type C progenitor toxin into HT-29 cells. *Biochem Biophys Res Commun* .**319**(2), 327–333.

Niwa, K., Koyama, K., Inoue, S., Suzuki, T., Hasegawa, K., Watanabe, T., Ikeda, T. and Ohyama, T. (2007). Role of nontoxic components of serotype D botulinum toxin complex in permeation through a Caco-2 cell monolayer, a model for intestinal epithelium. *FEMS Immunol Med Microbiol*. **49**:346–352.

Nölling, J., Breton, G., Omelchenko, M.V., Makarova, K.S., Zeng, Q., Gibson, R., Lee, H.M., Dubois, J., Qiu, D., Hitti, J., Wolf, Y.I., Tatusov, R.L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M.J., Bennett, G.N., Koonin, E.V., Smith, D.R. (2001). Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J. Bacteriol.* **183**, 4823-4838.

Oguma, K., Fujinaga, Y. and Inoue, K. (1997). *Clostridium botulinum* toxin. *J Toxicol Toxin Rev*. **16**, 253–266.

Oguma, K., Inoue, K., Fujinaga, Y., Yokota, K., Watanabe, T., Ohyama, T., Takeshi, K. and Inoue, K. (1999). Structure and function of *Clostridium botulinum* toxins. *Microbiol Immunol*. **39**(3), 161-8.

Ohishi, I., Sugii, S. and Sakaguchi, G. (1977). Oral toxicities of *Clostridium botulinum* toxins in response to molecular size. *Infect Immun*. **16**(1), 107–109.

Palma, L., Muñoz, D., Berry, C., Murillo, J. and Caballero, P. (2014). *Bacillus thuringiensis* Toxins: An Overview of Their Biocidal Activity. *Toxins*. **6**: 3296-3325.

Pérez-García, G., Basurto-Ríos, R. and Ibarra,J. E. (2010). Potential effect of a putative σH-driven promoter on the over expression of the Cry1Ac toxin of *Bacillus thuringiensis*. *J.Invert. Path.* **104**(2): 140-146.

Pigott C. R. and Ellar, D.J. (2007). Role of receptors in *Bacillus thuringiensis* crystal toxin activity. *Microbiol. Mol. Biol. Rev.* **71** (2): 255-281.

Piuri, M., Sanchez-Rivas, C. and Ruzal, S.M. (1998). A novel antimicrobial activity of a *Paenibacillus polymyxa* strain isolated from regional fermented sausages. *Lett Appl Microbiol*. **27**:9-13.

Poulet, S., Hauser, D., Quanz, M., Niemann, H. and Popoff, M. R. (1992). Sequences of the botulinal neurotoxin E derived from *Clostridium botulinum* type E (strain Beluga) and *Clostridium butyricum* (strains ATCC 43181 and ATCC 43755). *Biochem Biophys Res Commun.* **183**(1), 107-13.

Porter, A. G., Davidson, E. W. and Liu, J.-W. (1993). Mosquitocidal toxins of bacilli and their genetic manipulation for effective biological control of mosquitoes. *Microbiol. Rev.* **57:** 838–861.

Rao, D. R., Mani, T. R., Rajendran, R., Joseph, A. S., Gajanana, A., and Reuben, R. (1995). Development of a high level of resistance to *Bacillus sphaericus* in a field population of *Culex quinquefasciatus* from Kochi, India. *J Am Mosq Control Assoc.* **11**: 1-5.

Rasko, D. A., Altherr, M. R., Han, C. S. and Ravel, J. (2005). Genomics of the Bacillus cereus group of organisms. FEMS Microbiol Rev 29: 303-329.

Rummel, A and Binz, T. (2013). Botulinum Neurotoxins. *Current Topics in Microbiology and Immunology* **364**, DOI: 10.1007/978-3-642-33570-9_1, *Springer-Verlag Berlin Heidelberg*.

Saengwiman, S., Aroonkesorn, A., Dedvisitsakul, P., Sakdee, S., Leetachewa, S., Angsuthanasombat, C. and Pootanakit, K. (2011). *In vivo* identification of *Bacillus thuringiensis* Cry4Ba toxin receptors by RNA interference knockdown of glycosylphosphatidylinositol-linked aminopeptidase N transcripts in *Aedes aegypti* larvae. *Biochem Biophys Res Commun* **407**: 708-713.

Sankaranarayanan, R., K. Sekar, R. Banerjee, V. Sharma, A. Surolia, and M. Vijayan. (1996). A novel mode of carbohydrate recognition in jacalin, a Moraceae plant lectin with a β-prism fold. *Nat. Struct. Biol.* **3**:596-603.

Schiavo, G., Benfenati, F., Poulain, B., Rossetto, O., Polverino de Laureto, P., DasGupta, B.R. and Montecucco, C. (1992) .Tetanus and botulinum B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin. *Nature*. **359**, 832-835.

Schiavo, G., O. Rossetto, A. Santucci, B. R. DasGupta and C. Montecucco (1992). Botulinum neurotoxins are zinc proteins. *J Biol Chem*. **267**(33), 23479-83.

Schiavo, G., Matteoli, M. and Montecucco, C. (2000). Neurotoxin affecting neuroexocytosis. *Physiol Rev*. **80**, 717– 766.

Schiavo, G., Santucci, A., Dasgupta, B. R., Mehta, P. P., Jontes, J., Benfenati, F., Wilson, M. C. and Montecucco, C. (1993a). Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct COOH terminal peptide bonds. *FEBS Lett.* **335**:99–103.

Schiavo, G., Shone, C. C., Rossetto, O., Alexander, F. C. and Montecucco, C. (1993b). Botulinum neurotoxin serotype F is a zinc endopeptidase specific for VAMP/synaptobrevin. *J Biol Chem.* **268**:11516–11519.

Schiavo, G. and van der Goot, F. G. (2001). The bacterial toxin toolkit. *Nat Rev Mol Cell Biol*. **2**:530–537.

Schnepf, E., N. Crickmore, J. VanRie, D. Lereclus, J. Baum, J. Feitelson, D. R. Zeigler, and D. H. Dean. (1998). *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev*. **62**:775-806.

Seleena, P. and Lee, H. P. (1994). Absence of plasmid in mosquitocidal *Clostridium bifermentans* serovar *malaysia*. *Southeast Asian J. Trop. Med. Public Health* **25(2):** 394-396.

Seleena, P., Lee, H. L. and Lecadet, M. M. (1997). A novel insecticidal serotype of *Clostridium bifermentans*. *J Am Mosq Control Assoc.* **3(4):** 395-7.

Sharma, S.K., Ramzan, M.A. and Singh, B.R. (2003). Separation of the components of type A botulinum neurotoxin complex by electrophoresis. *Toxicon*. **41**, 321-331.

Smith, T. J., Lou,J., Geren, I.N., Forsyth, C. M., Tsai, R., Laporte, S. L., Tepp, W. H., Bradshaw, M., Johnson, E. A., Smith, L. A. and Marks, J. D. (2005). Sequence variation within botulinum neurotoxin serotypes impacts antibody binding and neutralization. *Infect Immun*. **73**(9), 5450-7.

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**: 1337-1349.

Storer, N.P., Babcock, J.M., Schlenz, M., Meade, T., Thompson, G.D., Bing, J.W. and Huckaba, R.M. (2010). Discovery and characterization of field resistance to Bt Maize: *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in Puerto Rico. *J Econ Entomol* **103**: 1031–1038.

Strotmeier, J., Lee, K., Völker, A. K., Mahrhold, S., Zong, Y., Zeiser, J., Zhou, J., Pich, A., Bigalke, H., Binz, T., Rummel, A. and Jin, R. (2010). Botulinum neurotoxin serotype D attacks neurons via two carbohydrate-binding sites in a ganglioside-dependent manner. *Biochem J*. **431**(2):207–216.

Strotmeier, J., Gu, S., Jutzi, S., Mahrhold, S., Zhou, J., Pich, A., Eichner, T., Bigalke, H., Rummel, A., Jin, R. and Binz, T. (2011). The biological activity of botulinum neurotoxin type C is dependent upon novel types of ganglioside binding sites. *Mol Microbiol.* **81**(1):143–156.

Sutton, R.B., Fasshauer, D., Jahn, R.and Brünger, A.T. (1998). Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature*. **395**, 347-353.

Sugawara, Y., Matsumura, T., Takegahara, Y., Jin, Y., Tsukasaki, Y., Takeichi, M. and Fujinaga, Y. (2010). Botulinum hemagglutinin disrupts the intercellular epithelial barrier by directly binding E-cadherin. *J. Cell Biol.* **189**(4): 691-700.

Swaminathan, S. and Eswaramoorthy, S. (2000). Crystallization and preliminary X-ray analysis of *Clostridium botulinum* neurotoxin type B. *Acta Crystallogr D Biol Crystallogr.* **56**, 1024-6.

Tonozuka, T., Oguma, K. and Nishikawa, A. (2014). Traffic of Botulinum Toxin Complex: Crystal Structure and Role of Toxin Complex. In *Glycoscience: Biology and Medicine.* pp 955-958.

Tabashnik, B.E., Gassman, A.J., Crowdwer, D.W. and Carriere, Y. (2008). Insect resistance to Bt crops: evidence versus theory. *Nat Biotechnol* **26**: 199–202.

Thie´ry, I., Hamon, S. V., Dumanoir, C. and de BarjaC, H. (1992a). Vertebrate safety of *Clostridium bifermentans* serovar *malaysia*, a new larvicidal agent for vector control. *J. Econ. Entomol.* **85:** 1618–1623.

Thie´ry, I., Hamon, S., Gaven, B. and de BarjaC, H. (1992b). Host range of *Clostridium bifermentans* serovar *malaysia*, a mosquitocidal anaerobic bacterium. *J. Am. Mosq. Control Assoc.* **8:**272–277.

Thompson, A.A., Jiao, G.S., Kim, S., Thai, A., Cregar-Hernandez, L., Margosiak, S.A., Johnson, A.T., Han, G.W., O'Malley, S. and Stevens, R.C. (2011). Structural characterization of three novel hydroxamate-based zinc chelating inhibitors of the *Clostridium botulinum* serotype A neurotoxin light chain metalloprotease reveals a compact binding site resulting from 60/70 loop flexibility. *Biochemistry*. **50**, 4019-4028

Turton, K., Chaddock, J. A. and Acharya, K. R. (2002). Botulinum and tetanus neurotoxins: structure, function and therapeutic utility. *Trends Biochem Sci* .**27**, 552-8.

Uotsu, N., Nishikawa, A., Watanabe, T., Ohyama, T., Tonozuka, T., Sakano, Y. and Oguma, K. (2006). Cell internalization and traffic pathway of *Clostridium botulinum* type C neurotoxin in HT-29 cells. *Biochim Biophys Acta* .**1763**(1), 120–128.

van Rensburg, J. B. J. (2007). First report of field resistance by stem borer *Busseola fusca* (Fuller) to Bt-transgenic maize. *S Afr J Plant Soil*. **24**:147–151.

Whelan, S. M., Elmore, M. J., Bodsworth, N. J., Brehm, J. K., Atkinson, T. and Minton, N. P. (1992). Molecular cloning of the *Clostridium botulinum* structural gene encoding the type B neurotoxin and determination of its entire nucleotide sequence. *Appl Environ Microbiol*. **58**(8), 2345-54.

WHO (2012). G;obal strategy for Dendue prevention and control 2012-2020. World Health Organization, Geneva, Switzerland.

WHO (2012). World malaria report 2012. World Health Organization, Geneva, Switzerland.

Wirth, M. C., Delecluse, A., Federici, B. A. and Walton, W. E. (1998). Variable cross-resistance to Cry11B from Bacillus thuringiensis subsp. jegathesan in Culex quinquefasciatus (Diptera: Culicidae) resistant to single or multiple toxins of Bacillus thuringiensis subsp. israelensis. Appl Environ Microbiol 64: 4174-4179.

Wirth, M. C., Walton, W. E. and Federici, B. A. (2010). Evolution of resistance to the Bacillus sphaericus Bin toxin is phenotypically masked by combination with the mosquitocidal proteins of Bacillus thuringiensis subspecies israelensis. Environ Microbiol 12: 1154-1160.

Yamashita, S., Yoshida, H., Uchiyama, N., Nakakita, Y., Nakakita, S., Tonozuka, T., Oguma, K., Nishikawa, A. and Kamitori, S. (2012). Carbohydrate recognition mechanism of HA70 from *Clostridium botulinum* deduced from X-ray structures in complexes with sialylated oligosaccharides. *FEBS Lett* . **586**(16), 2404–2410.

Yamasaki, S., Baumeister, A., Binz, T., Blasi, J., Link, E., Cornille, F., Roques, B., Fykse, E. M., Sudhof, T. C. and Jahn, R. (1994) Cleavage of members of the synaptobrevin/VAMP family by types D and F botulinal neurotoxins and tetanus toxin. *J Biol Chem.* **269**:12764–12772.

Yang, Y., Chen, H., Wu, Y., Yang, Y. and Wu, S. (2007). Mutated cadherin alleles from a field population of Helicoverpa armigera confer resistance to Bacillus thuringiensis toxin Cry1Ac. Appl Environ Microbiol 73: 6939-6944.

Yiallouros, M., Storch, V., Thie´ry, I. and Becker, N. (1994). Efficacy of *Clostridium bifermentans* serovar *malaysia* on target and nontarget organisms. *J. Am. Mosq. Control Assoc.* **10:**51–55.

Young, M., Staudenbauer, W. L., and Minton, N. P. (1989). Genetics of Clostridium. In "Clostridia" (N. P. Minton and D. J. Clarke, Eds.), pp. 63–104. Plenum, New York. Young, M., Minton, N. P., and Staudenbauer, W. L. (1989). Recent advances in the genetics of the clostridia. *FEMS Microbiol. Rev*. **63**, 301–326.

Zhang, X., Candas, M., Griko, N. B., Taussig, R. and Bulla, L. A. J. (2006). A mechanism of cell death involving an adenylyl cyclase/PKA signaling pathway is induced by the Cry1Ab toxin of *Bacillus thuringiensis*. *Proc Natl Acad Sci* **103**: 9897-9902.

Zhang, R., Hua, G., Andacht, T. M. and Adang, M. J. (2008). A 106-kDa aminopeptidase is a putative receptor for *Bacillus thuringiensis* Cry11Ba toxin in the mosquito *Anopheles gambiae*. *Biochemistry* **47**: 11263-11272.

Zhang, H., Yin,W., Zhao, J., Jin, L., Yang, Y., Wu, S., Tabashnik, B.E. and Wu, Y. (2011). Early warning of cotton bollworm resistance associated with intensive planting of Bt cotton in China. *PLoS ONE* **6**: e22874.
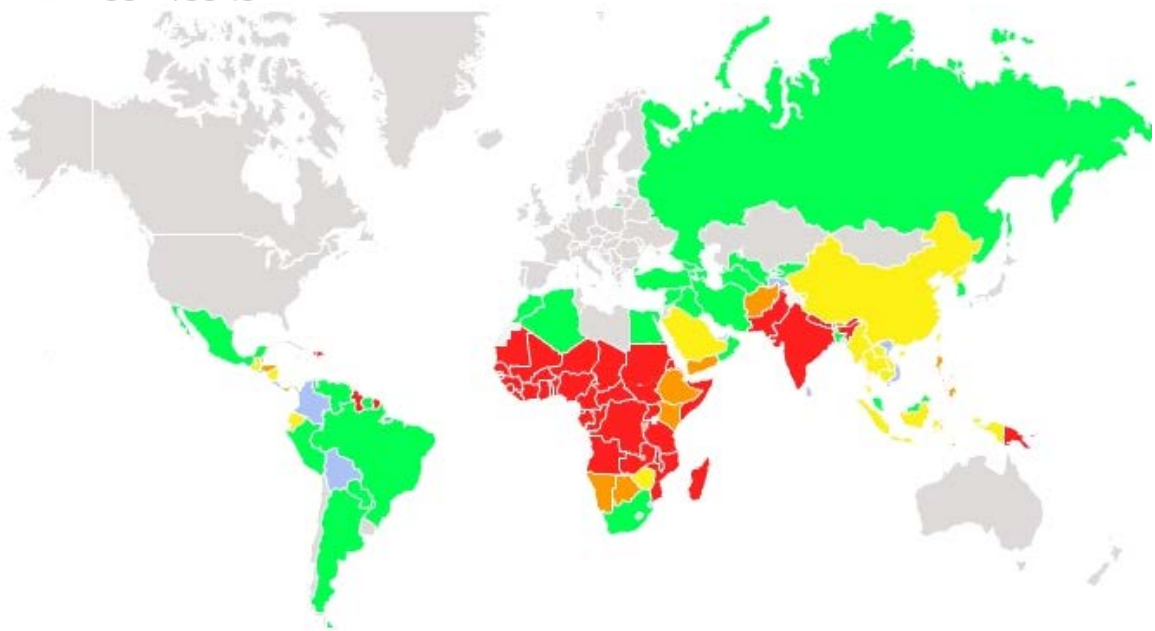
Figure 1.1: Percent of Population at risk of malaria. Population at risk (High+Low):

High=population living in areas (reported malaria incidence ≥ 1 per 1000/year) defined at

administravive level 2 or lower. Low=population living in areas (reported malaria

incidence < 1 per 1000/year). Customized map created using Global Malaria Mapper

(http://www.who.int/malaria/publications/world_malaria_report/global_malaria_mapper/

en/) using data from the *World Malaria Report 2012*.
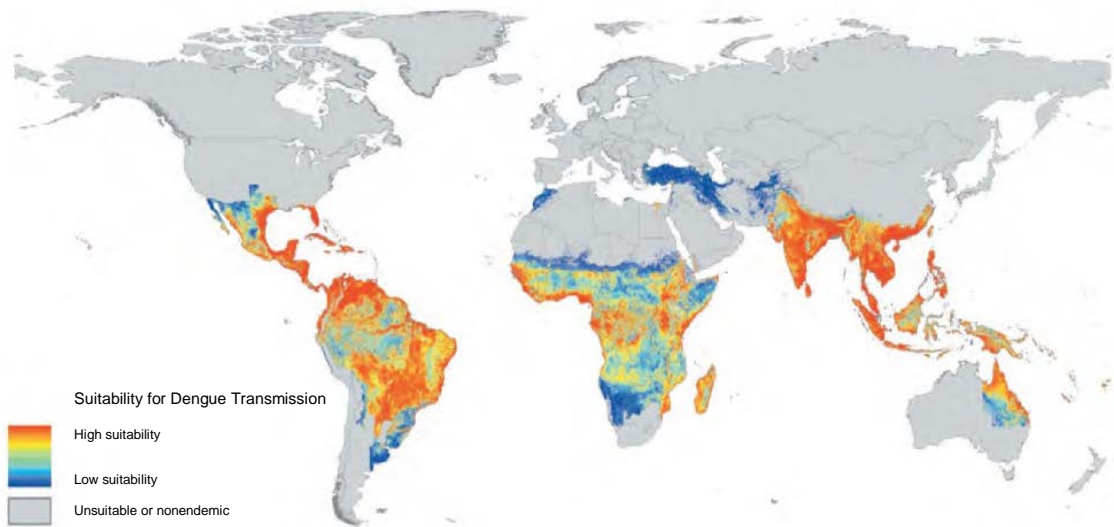
Figure 1.2. Distribution of global dengue risk. Dengue fever occurs in subtropical and

tropical regions in the world. Data from WHO, 2012.
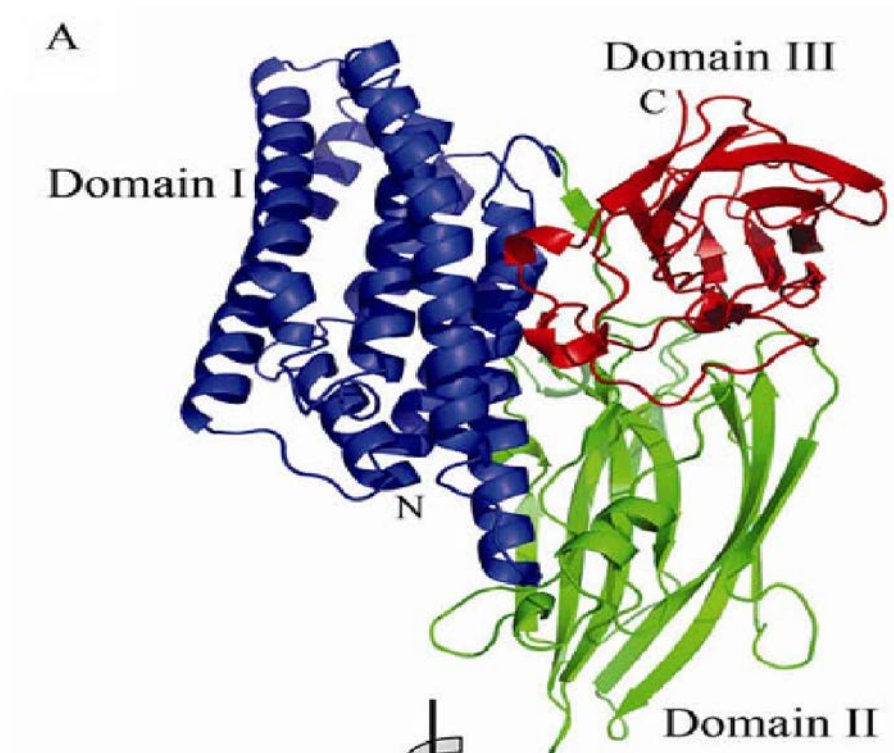
Figure 1.3: 3-D crystal structure of Cry8Ea determined at 2.20Å (PBD code: 3EB7). The three domains of the protein are represented with different colours with domain I coloured blue, domain II coloured green while domain III is coloured red (Taken from Guo et al., 2009).
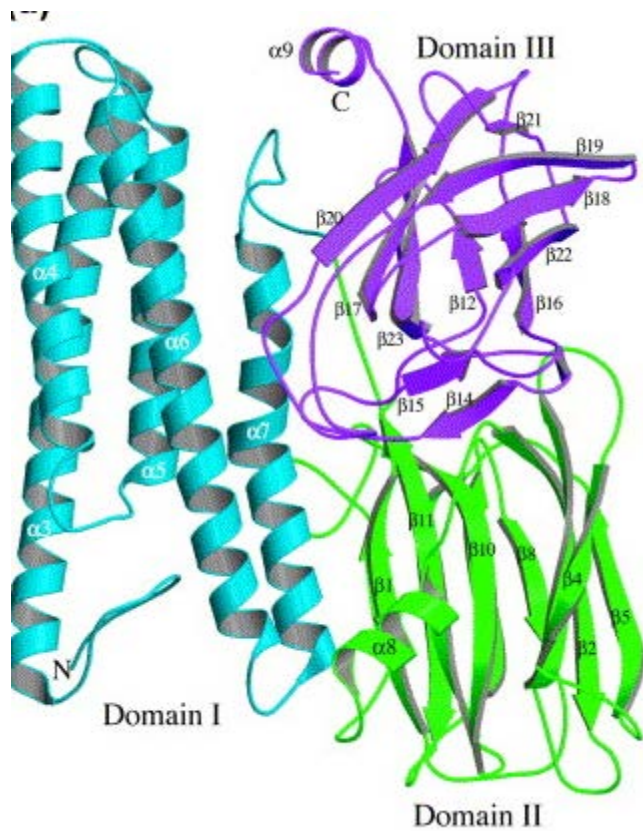
Figure 1.4: 3-D crystal structure of Cry4Ba determined at 1.75Å (PBD code: 1W99).
The three domains of the protein are represented with different colours with domain I
colored blue, domain II colored green while domain III is colored purple (Taken from
Boonserm et al., 2005).
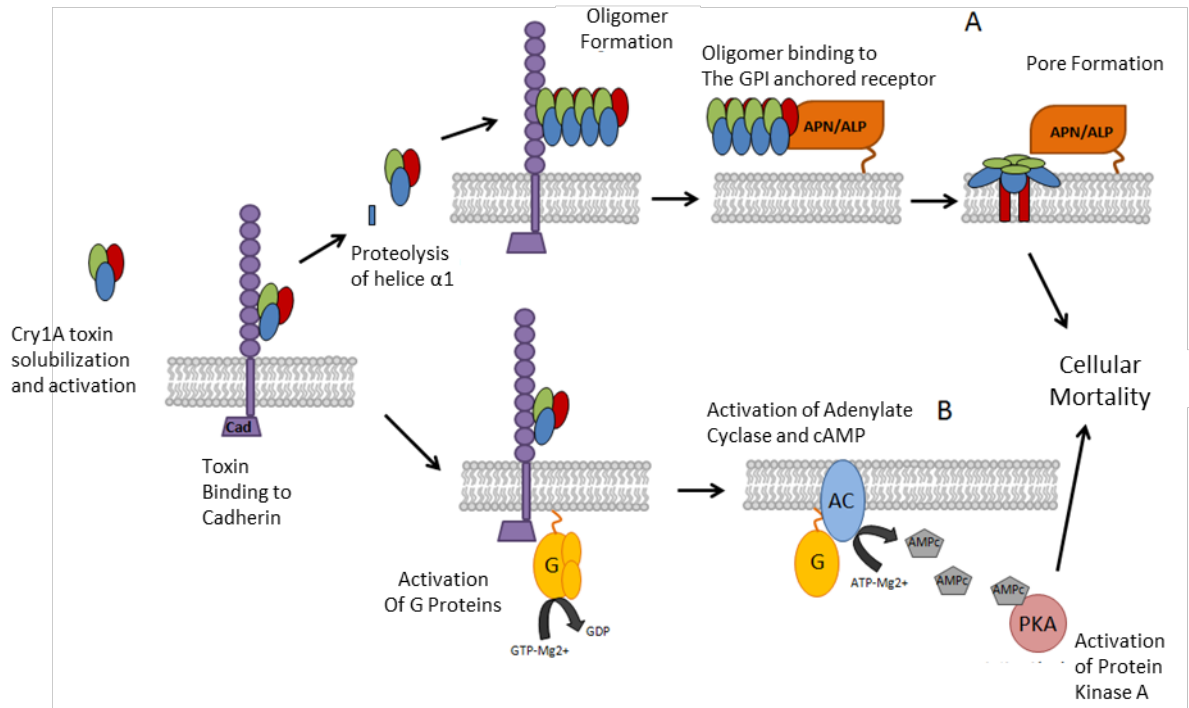
Figure 1.5: Mechanism of action of Cry1 toxins in Lepidoptera. (A) Bravo (pore formation) model. (B) Zhang (signal transduction) model. Taken from Soberon et al., 2009.
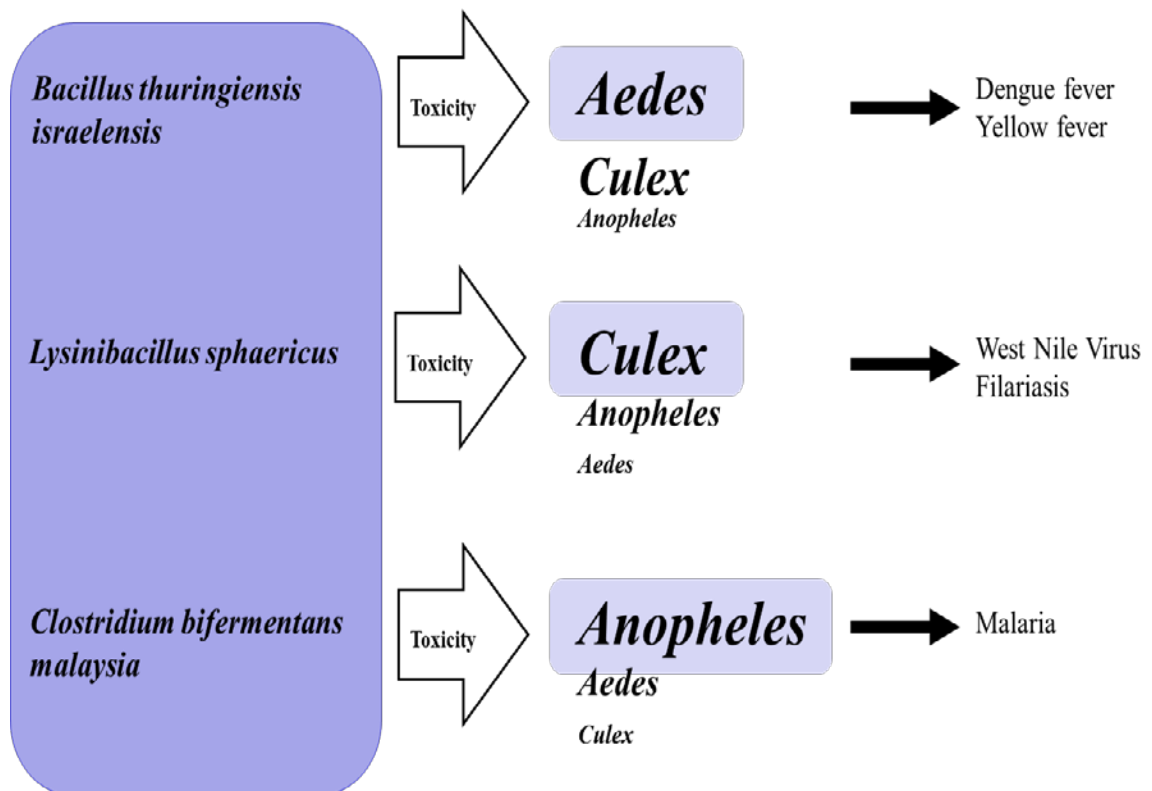
Figure1.6: Species specificity of various biological control agents. Bti shows highest toxicity against *Aedes* (major vector of dengue and yellow fevers) followed by *Culex* and *Anopheles.* Ls shows highest toxicity against *Culex* (major vector of west nile and filariasis) followed by *Anopheles* and *Aedes.* But Cbm is the only one that shows highest toxicity against *Anopheles* (major vector of malaria) followed by *Aedes* and *Culex.*

BoNT complex

12S toxin (M toxin) 16S toxin (L toxin) 19S toxin (LL toxin)

BoNT (Neurotoxin) [(L chain) (H chain)]

NTNHA (non-toxic non-HA)

HA (Hemagglutinin) [HA1 HA2 HA3]

Non-toxic neurotoxin-associated proteins (NAPs)

Toxin type	12S toxin	16S toxin	19S toxin
A (A1)	+	+	+
B, C, D	+	+	-
E, F, A2, A3	+	-	-
G	-	+	-
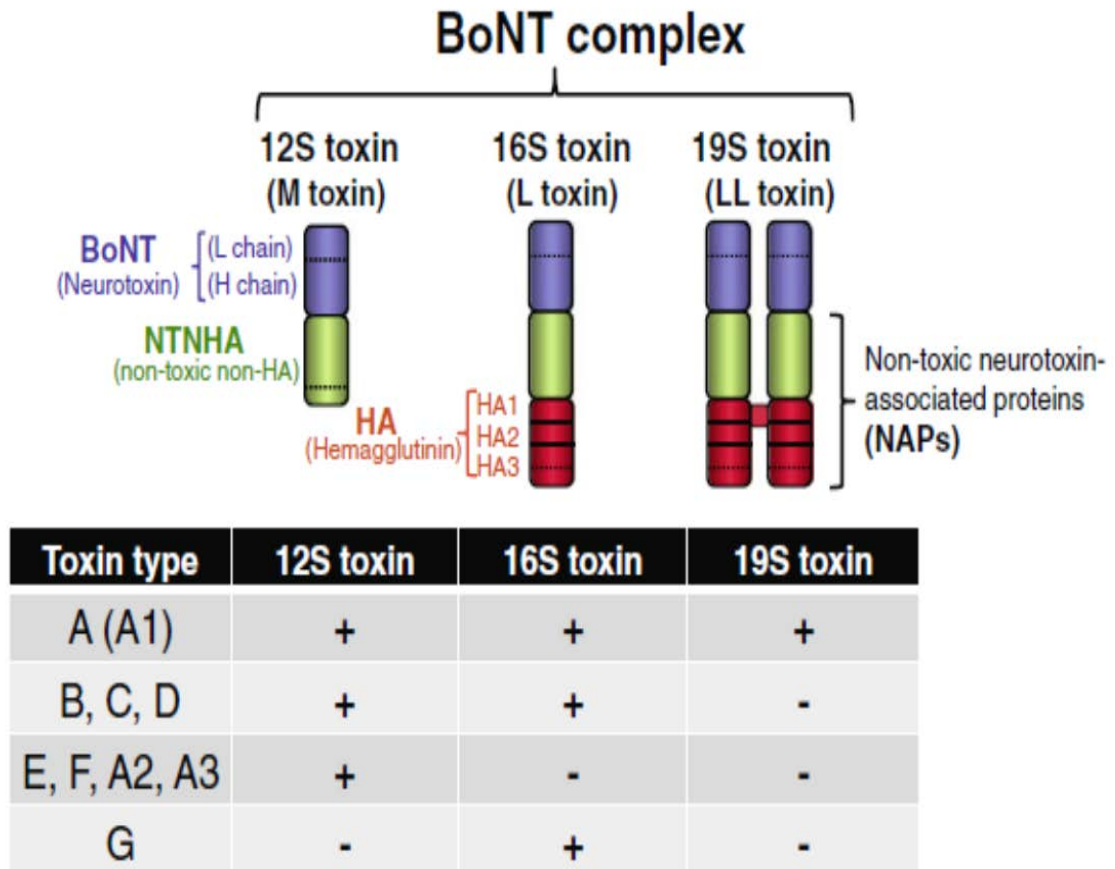
Fig 1.7: Schematic representation of botulinum neurotoxin BoNT complexes. Three major forms of BoNT complexes, 12S toxin (M toxin), 16S toxin (L toxin), and 19S toxin (LL toxin), are presented. The 12S progenitor toxin complex consists of a single BoNT molecule and one NTNH protein, but lacks the associated proteins responsible for hemagglutination activity. The 16S progenitor toxin, in addition to the components found in the 12S complex, contains three hemagglutinin (HA) proteins, HA70, HA33, and HA17. The secreted and most toxic form 19S complex is believed to be a dimer of two 16S toxins linked by an additional HA33 protein. Lower panel BoNT/A1 is produced by C. botulinum in three forms: 12, 16, and 19S toxins. BoNT/B, C, and D are produced in two forms: 12S and 16S. BoNT/A2, A3, E, and F are produced as 12S. BoNT/G complex is produced as 16S, but there is a lack of HA1 gene in the BoNT/G-cluster of genes. + presence of the complex forms; - not present (Taken from Fujinaga et al. 2013. Curr Topic Microbiol. Immunol. 364).
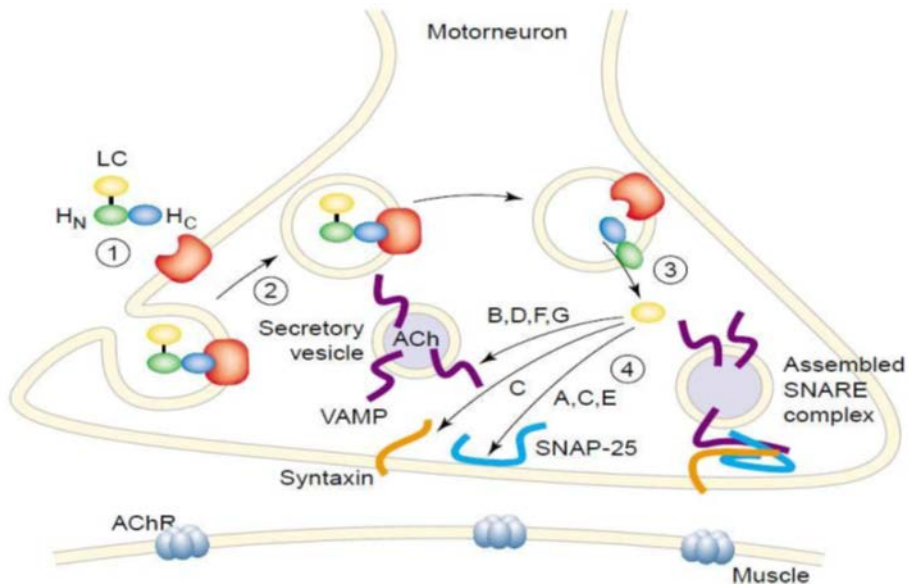
Figure 1.8: Mode of action of botulinum neurotoxins The schematic shows a four-step mechanism, (1) cell binding of the BoNT, (2) endocytosis, (3) translocation of LC in the cytosol, (4) LC cleavage of one of the SNARE proteins. Taken from Turton et al., 2002.
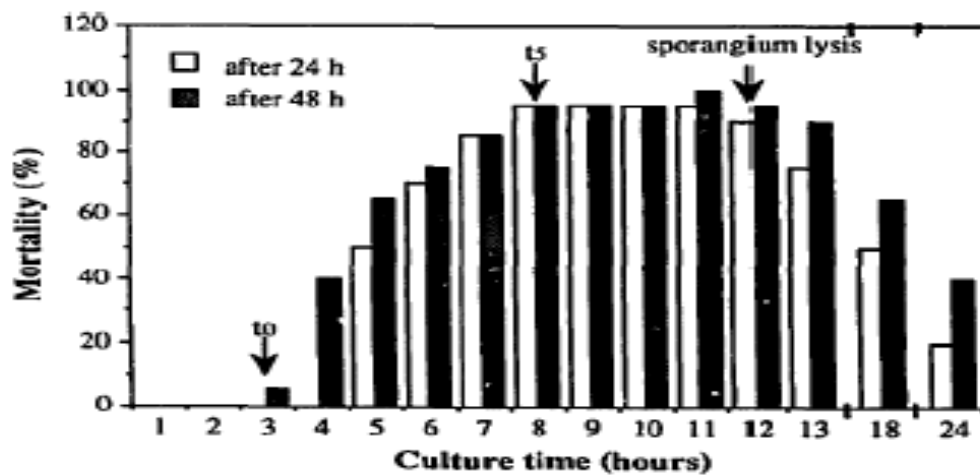
Figure 1.9: Mosquitocidal activity of the sporulating cultures of *C. bifermentans* serovar *malaysia*. Mortality is expressed as percent of dead 2nd instar *A. stephensi* larvae after 24 and 48 h exposure. (Figure from Charles et al., 1990)

A.

B.
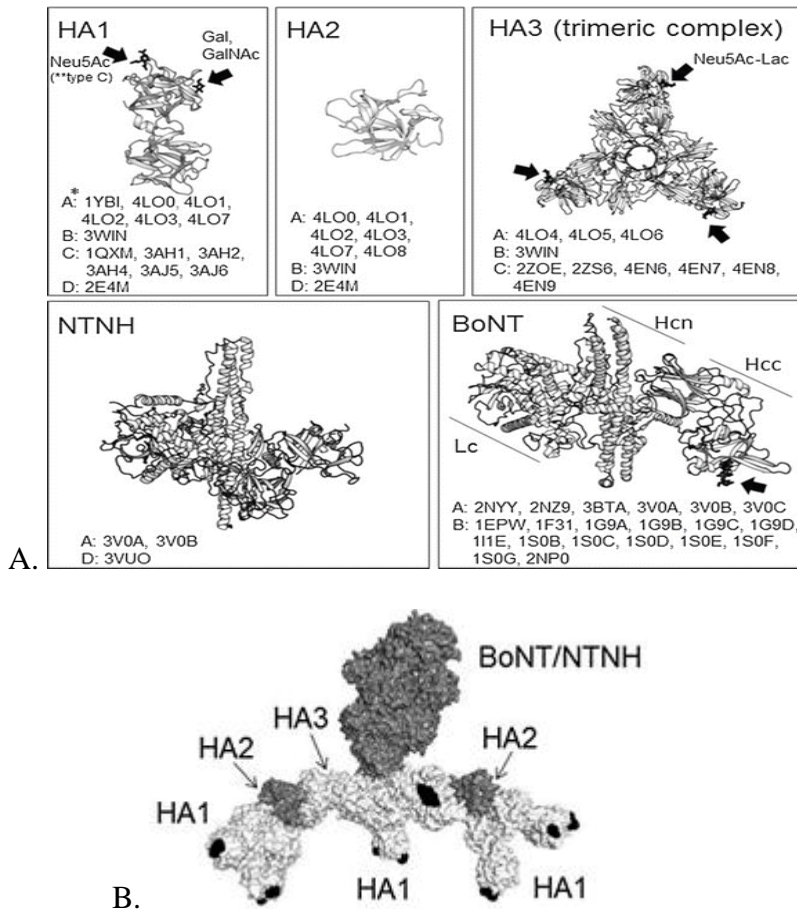
Figure 1.10: The crystal structures of 16S BoNT and its nontoxic components.

A. The carbohydrate-binding sites provided by X-ray crystallography are indicated by the arrow. *: Serotypes and PDB codes. HA1 from types A and B has been predicted to have a single carbohydrate-binding site, while type C HA33 has two site indicated by ** (Nakamura et al. 2009).

B. The whole structure of 16S toxin was estimated by X-ray crystallography and electron microscopy (Benefield et al. 2013; Lee et al. 2013; Amatsu et al. 2013). Carbohydrate-binding sites are shown in black. The 16S toxin has at least nine carbohydrate-binding sites and forms a large tripod-shaped complex. The 16S toxin is a complex with BoNT : NTNH : HA1 : HA2 : HA3 stoichiometry of 1:1:6:3:3

(Taken from Tonozuka et al., 2014)

55

Chapter 2

Genome Structure of *Clostridium bifermentans* subsp. *malaysia*

Abstract

Biological control of mosquitoes and blackflies, that transmit severe human diseases, have relied heavily on the use of *Bacillus thuringiensis* subsp. *israelensis* and *B. sphaericus* as larvicidal agents. The primary threat to the long-term efficacy of Bt toxins is the evolution of resistance by pests already shown in the laboratory and to Bt sprays in the field. Knowledge of the complete genome sequence of *Clostridium bifermentans* serovar *malaysia* is expected to facilitate the development of novel biological control solutions with different mode of actions and successful alternative to Bt as a useful IPM organism. The genome sequence of the only known anaerobic mosquitocidal bacterium *Clostridium bifermentans* serovar *malaysia* (Cbm) has been determined by the next generation sequencing approach. The genome consists of a 3,723,765 bp chromosome encoding 3,835 ORFs including the previously implicated larvicidal Cry genes. Additional virulence-related factors could be identified for further work.

Introduction

Clostridium is a gram positive, obligate anaerobic, spore-forming anaerobe with a rod shaped cellular morphology and a pilus-based motility, forming heat-stable endospores (Rummel and Binz, 2013). The genus comprises of numerous pathogenic and non-pathogenic strains. Among the non-pathogenic strains studied are Clostridia with unique metabolic functions. The pathogenic strains are responsible for severe diseases in humans and animals because of production of potent protein toxins. Of the more than 150 *Clostridium* species, 35 are considered pathogenic and 15 of those produce known protein toxins (Binz and Rummel, 2013). The most common strains among the toxigenic *Clostridium* are *C. diffile*, *C. tetani*, *C. perfringens*, and *C. botulinum*. *C. difficile* causes pseudomembranous colitis, a severe infection of the colon, characterized by diarrhea, fever, and abdominal pain. *C. tetani* causes continuous muscle contractions, commonly known as lock-jaw. *C. perfringens* is the causative agent of gas gangrene and necrotic enteritis. *C. botulinum* is the causative agent of botulism, a nerve paralysis disease.

2.1 Clostridial Genomes

As typical of the Firmicutes, Clostridial genomes maintain an exceptionally low G+C content of ~28% (Myers et. al., 2006; Shimizu et. al., 2002). The genomes of several clostridial species and their strains have been sequenced, revealing an average ~3.0 – 7.0 Mbp genome size, showing a considerable amount of variation. These genomic differences cluster to over 300 different genomic islands and can encode genes involved with metabolism, capsule synthesis, toxin production, and mobile genetic

elements. In addition to its chromosome, *Clostridium sp.* can harbor several large

plasmids, many of which contain toxins and other virulence factors vital for the

establishment and maintenance of infections (Miyamoto et. al., 2006; Rood, 1998). Based

on the relative content of G+C of the chromosome, *Clostridium* species can be classified

into two groups: the extremely low G+C group (ranging between 24% and 31% G+C

contents) includes *C. perfringens*, and *C. acetobutylicum*, and the second group, showing

slightly higher G+C contents (ranging between 32 and 40%) includes thermophiles such

as *C. thermocellum* and *C. cellolyticum* (Young et al., 1989). AT-rich stretches seem to

cover mainly non-coding regions. The chromosomes contain 3000 – 4000 open reading

frames and encode 10 – 14 ribosomal RNAs. Due to their exceptionally low G+C

contents, Clostridia have particular codon usage. In *C. acetobutylicum* and *C. perfringens*

arginine is encoded by AGA in 93% cases, while the codons CGC, CGA and CGG are

not used. Clostridia generally use AUG (Met), GUG (Val) and UUG (Leu) as

translational start codons. Particularly striking are the ribosome binding sites (RBS) of

Clostridia that contain preferentially G residues; the most redundant RBS encountered

being GGAGG and GGGGG. Currently in NCBI Database, there are sequences of 38

clostridial genomes in various stages of progress and completion, some of which are

listed in Figure 2.1.

2.2 Genome of Bti, a successful biocontrol agent

   Bacillus thuringiensis isolates are the biological control agents that are most

widely used to eradicate insect pests of crops or vectors of human disease. For the last

decade, *Bacillus thuringiensis* subsp. *israelensis* is the bio insecticide of choice in programs worldwide to control mosquitoes and blackfly vectors (Berry et al., 2002). The genome of *B. thuringiensis* strains range from 2400 to 5700 Mb present as chromosomal and extrachromosomal elements (Carlson et al., 1994, 1996; Crickmore et al., 2014). In the past decades, more than 700 cry gene sequences that code for crystal (Cry) proteins have been identified (Crickmore et al., 2014; Van Frankenhuyzen, 2009) and large plasmids appear to be the usual location for these genes (Kronstad et al., 1983; Ward et al., 1983), although chromosomal homologs have also been identified (Carlson et al., 1993). Plasmids in *B. thuringiensis* strains range from 4.56 to 228 kb (Gonzalez et al., 1980; Lereclus et al., 1982; Baum et al., 1992). *B. thuringiensis* harbors a plethora of transposable elements including insertion sequences and transposons, often located in the vicinity of the toxin gene, thus facilitating intra/ inter molecular genome mobility (Battisti et al., 1985; Green et al., 1989). For instance, the insect pathogenicity of *Bacillus thuringiensis* subsp. *israelensis* depends on the presence of the pBtoxis megaplasmid (Faust et al., 1983) that encodes all six of the previously described toxins in this isolate (Cry4Aa, Cry4Ba, Cry10Aa, Cry11Aa, Cyt1Aa, and Cyt2Ba)(Ben-Dov et al., 1999; Berry et al., 2002).

By unraveling the gene content of a genome and the corresponding metabolic details, there is the possibility that new doors may open and lead to the development of innovative strategies to modify bacteria for applications such as biological control agents of pests.

2.3 Next Generation Sequencing Technologies

Until 2005, most DNA sequencing reads were generated by Sanger technology, published by Dr. Fred Sanger in 1977(Sanger et al., 1977). Sanger sequencing utilizes dideoxynucleotides (ddNTP's) that prevents the addition of more nucleotides on a DNA strand. Adding fluorescently or radioactively labeled ddNTP's to template DNA will produce a pool of DNA molecules in which some molecules terminate at every position along the target DNA.The target DNA sequence is determined by reading the labeled ddNTP at each position by using capillary gel electrophoresis (Kim et al., 2008).

The 454 Genome Sequencer (GS) platform is based the pyrosequencing process (Figure 2.2) in which one nucleotide at a time is washed over several copies of the sequence to be determined, causing polymerases to incorporate the nucleotide if it is complementary to the template strand. The incorporation stops if the longest possible stretch of complementary nucleotides has been synthesized by the polymerase. In the process of incorporation, one pyrophosphate per nucleotide is released and converted to adenosine triphosphate (ATP) by an ATP sulfurylase. The ATP drives the light reaction of luciferases present and the emitted light signal is measured.

The reversible terminator technology used by the Illumina Genome Analyzer (Bentley et al., 2008; Turcatti et al., 2008) (Figure 2.3) requires two different adapters are added to the 5' and 3' ends of all molecules. The double-stranded library is melted to obtain single stranded DNAs and pumped through the channels of a flow cell. These oligonucleotides will hybridize to the single stranded library molecules. These identically

oriented copies of the same sequence can be sequenced by hybridizing the sequencing primer onto the adapter sequences.

Pacific Bioscience extracts high-resolution data from individual chromosomes using Single Molecule Real Time (Figure 2.4) technology (Korlach et al., 2008). This technology performs the sequencing reaction on silicon dioxide chips with a 100nm metal containing thousands of tens-of nanometers diameter holes, so called zero-mode waveguides (ZMWs). A single DNA polymerase is fixed to the bottom of the surface within each ZMW. Nucleotides with different fluorescent dyes attached to the phosphate chain are added and DNA molecules are sequenced by measuring the fluorescent signal of nucleotides being incorporated by DNA polymerase (Eid et al., 2009).

Optical mapping (Zhou et al., 2007) is a technique for constructing ordered, genome-wide, high-resolution restriction maps from single, stained molecules of DNA. By mapping the location of restriction enzyme sites along the unknown DNA of an organism, the spectrum of resulting DNA fragments collectively serve as a unique fragment pattern for that sequence.

In the present study, we determined the draft genomes of *Clostridium bifermentans* and its two subspecies –*malaysia* and –*paraiba* that are mosquitocidal. Mosquitoes transmit some of the world's most life threatening and debilitating parasitic and viral diseases, including malaria (*Anopheles*), filariasis (*Culex*, *Mansonia* and some *Anopheles* spp) and dengue and yellow fevers (*Aedes aegypti*). *Clostridium bifermentans* serovar *malaysia*, (Cbm) is a gram positive spore forming strain isolated from Malaysia, has high toxicity to a number of *Anopheles*, followed by *Aedes* and *Culex* (Thiery et al.,

1992b) but is also not toxic to vertebrate (Thiery et al., 1992a). Genetic Analysis of putative insecticidal factors (Nicolas et al., 1993a; Nicolas et al., 1993b) led to identification of four genes. These genes were cloned and initially named *Cbm71* and *Cbm72* (Barloy et al., 1996) and two hemolysin-like genes, *Cbm17.1* and *Cbm17.2* (Barloy et al., 1998a). *Cbm71* and *Cbm72* were structurally related to δ-endotoxin genes of *Bacillus thuringiensis* and encoded proteins of 71.13kDa and 71.73kDa respectively, and are now named *Cry16Aa* and *Cry17Aa* (Barloy et al., 1996, 1998; Crickmore et al., 2011). *Cbm17.1* and *Cbm17.2* encoded for proteins of sizes 17.19kDa and 17.45kDa respectively and two proteins hare 44.6% similarity with the hemolysin of *Aspergillus fumigatus* but preliminary tests did not show hemolytic activity (Barloy et al., 1996, 1998). *Cry16Aa* and *Cry17Aa* were tested for toxicity and showed negligible activity, if any. Later, Juarez-Perez et al., 2001 showed that recombinant proteins produced by these cloned genes were not mosquitocidal, suggesting that the proteins responsible for mosquitocidal activity remain unidentified.

Since the classical cloning and protein methods did not work, our lab undertook a comparative genomic approach to identify larvicidal toxin genes from *C. bifermentans subsp malaysia*. We sequenced the entire genomes of *C. bifermentans* (Cb)*, C. bifermentans* subsp. *paraiba* (Cbp), *C. bifermentans subsp malaysia* (Cbm) and a *C. bifermentans* subsp. *malaysia* mutant (Cbm-mutant). The objective was to sequence the entire genomes of Cb, Cbm, Cbp and Cbm-mutant and then compare and analyze them to search for the presence and absence of known and putative genes, as well as for the regions of synteny and divergence. For this purpose, genomic DNA was extracted from

Cb, Cbm, Cbm-mutant and Cbp and sent for sequencing using Illumina sequencing platform and Roche's 454 sequencing platform. The sequence data was assembled de novo into contigs and scaffolds and annotated by BGI. Cultures of Cbm and Cbp were sent to OpGen where optical mapping of the genomes were performed. The optical map was used as a scaffold on which to align the contigs and scaffolds received from sequencing. The hypothesis is that if novel toxin genes are present then they will likely be plasmid-borne, but whole genome sequencing can be much easier than isolating and sequencing single plasmids, especially if the strain contains more than one plasmid.

Material and Methods

Bacterial strains and culture conditions

Clostridium bifermentans (Cb) (ATCC) was used as the wild type reference strain and *C. bifermentans* subsp *malaysia* (Cbm) and *C. bifermentans* subsp *paraiba* (Cbp) were from the collection of the Institute for Medical Research, Malaysia (Lee et al., 1990b). A *C. bifermentans* subsp *malaysia mutant* (Cbm-mutant) that lost its larvicidal activity was generated in my lab by Dr. Nadia Qureshi using gamma irradiation. Clostridial strains were streaked for isolation on TYG agar and liquid cultures were grown in TYG medium at 30°C under anaerobic conditions using BD GasPakEZ (Beckton-Dickinson Microbiology, Cockeysville, MD).

General DNA Techniques

All DNA manipulations were performed according to standard protocols. Primers used in this study were created using the sequences from existing databases. Genomic DNA was extracted from bacterial culture grown overnight at 30°C in TYG medium by using a Wizard DNA purification kit (Promega, Madison, WI) or DNeasy Blood and Tissue Kit (Qiagen, Germany) according to the manufacturer's instructions. Quantity and quality of the DNA were measured spectrophotometrically (Nanodrop ND-1000, NanoDrop Tech. Inc., Wilmington, DE). Polymerase chain reactions (PCR) were performed in an automated therma-cycler (C 1000 Touch ™, Bio Rad). Choice-Taq^{TM} mastermix DNA polymerase (Denville Scientific, Metuchen, NJ) was used for all PCR reactions for products below 2kb. Advantage 2x polymerase (Clonetech, Mountain View, CA) or Phusion polymerase (NEB) was used for all PCR reactions for products above 2kb. PCR products were separated in 1% agarose gels and subsequently cut and purified using Wizard SV Gel and PCR purification kits (Promega, Madison, WI). Sequencing of purified DNA products was performed by the Genomics Core facility at the University of California, Riverside.

Mutagenesis of Cbm with Cisplatin

Cbm cells were grown overnight in 10ml TYG media under anaerobic conditions at 30°C. An overnight Cbm culture (500 µl) was inoculated into 5ml BHI anaerobic tubes (Anaerobe Systems, CA) and grown for 2 hours anaerobically at 30°C. Cisplatin (2mg/ml) was then injected into the vials and the cultures were grown for 4 hours. Then

1µl, 10µl and 25µl aliquots of the 4 hour culture were plated on TYG plates and grown overnight at 30°C in anaerobic conditions. Rapid screen of colonies was done using five 3rd instar *Aedes* larvae in 1ml water and loss or decrease of mosquitocidal activity was noted. Those colonies that showed loss or decrease of toxicity were then grown in 10ml cultures as described before and bioassayed. In brief, 25 third-instar larvae were transferred to plastic cups containing 200 ml tap water and then fed whole culture of the Cbm-Cisplatin mutant colony at different concentration. Mortality was recorded at 24h.

Genome Sequencing and Mapping

Whole-genome sequencing was performed using three high-throughput sequencing techniques, Sanger sequencing, 454 pyrosequencing and Illumina sequencing platform. Single and paired-end reads for Cbm were performed at the Genomics Core, UC Riverside. Paired-end libraries (500bp and 5000bp) of Cb, Cbm, Cbm-mutant and Cbp were created and sequenced using Illumina sequencing platform at the Beijing Genomic Institute, China. Additionally, *de novo* whole genome sequencing of *C. bifermentans* subsp *malaysia* was performed using the PacBio SMRT system (Pacific Biosciences, Menlo Park, CA, USA) at National Centre for Genomic Research, Los Alamos, NM. *de novo* NcoI optical maps of the genomes of Cbm and Cbp were generated by OpGen technologies (Madison, WI). The optical maps were used as a guide scaffold on which to align the scaffolds received from sequencing.

Gap closure and Assembly validation

The Sanger sequences, 454 contigs, and BGI Cb Scaffold as well as BGI Cbp Scaffolds obtained from Illumina were used to confirm the assembly of BGI Cbm scaffolds. The BGI Cbm Illumina scaffolds were aligned against the listed sequences using BLASTn to confirm the orientations and integrity of the assembled sequences and to close gaps and link scaffolds together. Mauve™ (Darling et.al, 2004) was also used to align and compare the Cb, Cbm, Cbm-mutant and Cbp genomes scaffolds for consistency of Cbm sequencing assembly into scaffolds and any rearrangement patterns. PCR primers were designed and Sanger sequences of these PCR products were used to close the gaps between the Cbm scaffolds. *In silico* NcoI restriction maps of the NCGR scaffold using PacBio sequencing, BGI Cbm, BGI Cb and BGI Cbp scaffolds were constructed and aligned to the optical map of Cbm according to their restriction fragment pattern by using the Argus™ optical mapping system and analysed using MapSolver v.3.1 software (OpGen Technologies, Inc.). Final assembly was correlated with the optical map for further validation.

Sequence-to-map comparison

Comparisons between Optical maps and all the BGI, PacBio and 454 sequencing scaffolds were performed as follows. Sequence FASTA files were converted to *in silico* restriction maps via the MapViewer software (OpGen Technologies, Inc.) for direct comparison to the Optical maps. Comparisons were accomplished by aligning the sequence with the Optical maps according to their restriction fragment pattern.

Annotation of Scaffolds

Coding genes and predicted proteins were annotated by BGI against the NCBI non-redundant (nr) protein database (http://nCbi.nlm.nih.gov), the clusters of orthologous groups (COG) database, Swiss-Prot database, TrEMBL protein database and the Kyoto Encyclopedia of Genes and Genomes (KEGG) Genes and KEGG Pathway databases. Annotation of the re-assembled genome into two scaffolds was done with BLAST (Altschul et al., 1990) and and complemented with the RAST server (Aziz et al., 2008).

Comparative Analysis

Complete genome sequences of all the three *Clostridium* spp. were aligned and visualized in progressive mode using MAUVE. The sequences absent or present or rearranged in one or more genomes were further validated by PCR. *In silico* NcoI restriction maps of the three genomes scaffolds were constructed using the Argus™ optical mapping system and aligned to the optical map according to their restriction fragment pattern by using MapSolver v.3.1 software (OpGen Technologies, Inc.).

Additional Sequence Analysis

Candidate plasmids were identified by both plasmid preparations from overnight cultures using Wizard Plus SV Minipreps DNA purification kit (Promega, Madison, WI) and by visualization in MAUVE. Putative plasmids were confirmed by PCR and Sanger sequencing.

Nucleotide sequence accession number

The draft genome sequences of *C. bifermentans* subsp. *malaysia*, *C. bifermentans* subsp. paraiba, *C. bifermentans* subsp. *malaysia* mutant and *C. bifermentans* has not been deposited in any data base yet.

Results

Sequencing and general features of the *C. bifermentans* subsp *malaysia* genome.

Finding novel proteins is particularly important to manage the increasing resistance occurrence to Bt-based insecticides or Bt-plants reported for some species (Ferré et al., 2002.; Sayyed et al., 2000; Oppert et al, 1997). Today, next-generation sequencing (NGS) technologies (Shendure et al., 2008) provide a novel and useful tool for the discovery of completely novel insecticidal-toxin genes that would otherwise be difficult to identify. In this work, we used an efficient strategy of combining Illumina de novo sequencing with PacBio and 454 de novo sequencing, which allowed the genome to be sequenced completely and assembled correctly with just one gap remaining.

Genomic DNA of *C. bifermentans* subsp *malaysia, C. bifermentans* subsp *paraiba, C. bifermentans* and *C. bifermentans* subsp *malaysia* mutant strain was isolated and sent to BGI, China, for *de novo* whole genome sequencing using paired-end Illumina sequencing on 500bp and 5000bp libraries. Short reads were then assembled into genome sequence using SOAP de novo and gap filling and single base correction were performed via SOAPaligner. Scaffolds were the constructed using the mapping information from paired-end reads. BGI then sent us Cb, Cbp and Cbm assemblies which contained 39, 29

68

and 39 (Table 2.4) scaffolds respectively which collectively contained several hundred copies of transposons and ribosomal RNA coding regions. To get a more accurate and complete assembly, we also sent Cbm genomic DNA to NCGR where they created 10Kb libraries for us and sequenced them using Pacific Biosystems SMRT platform. NCGR sent us 185 consensus reads for "scaffolds".

Given this complexity, optical mapping was attempted to provide a structural backbone for aligning and orienting the scaffolds. Optical mapping permits assembly of whole-genome restriction endonuclease maps by digesting immobilized DNA molecules and determining the size and order of fragments (Zhou et al., 2007). Optical map of Cbm was produced using *NcoI* restriction enzyme. OpGen's Argus assembler program reconstructed the ordered restriction map of the genome. The Cbm optical map was aligned with in-silico genome assemblies of Cb, Cbp and Cbm generated from BGI Scaffold sequences, 454 sequences and PacBio Scaffolds. The map permitted alignment and orientation of only few scaffolds in each case (Figure 2.5, A, B, C and D). But collectively and along with gap closure using MAUVE (Figure. 2.11), they led to covering the whole genome with one gap.

Cbm genome assembly and Annotation

C. bifermentans subsp *malaysia* has a single, circular chromosome of approximately 3,723,765 bp with eight plasmids, with a largest plasmid of 109kb. Details of the general features of the Cbm genome are shown in Table 2.4. The G+C content of the chromosome averages 28.05%, which is similar to that of other *Clostridium* genomes.

The chromosome displays a clear GC skew transition typical of prokaryotic genomes, indicative of a bidirectional replication mechanism (Ravin et. al. 2003). GC skew analysis by Artemis Software and blast comparison were used to locate the origin of replication at the point with an excess of G over C corresponding to the beginning of the leading strand. The Cbm genome contains 3,835 putative CDSs with slightly higher GC content of 28.77% in the coding regions (Table 2.2). The genome shows a coding density of 85.05%, characteristic of most clostridial genomes (Fonknechten, et al., 2010; Zhou et al., 2014). Beijing Genome Institute sent us the alignment results with the following databases: KEGG, COG, SwissProt, TrEMBL, NR (Figures 2.6 and 2.7) in GFF format.

Comparison of chromosome organization of Cbm and Cbp

On comparing the optical maps of Cbm and Cbp, it was seen that the Cbp genome was smaller than Cbm genome and there were some regions in Cbm that showed no sequence similarity in Cbp. They were termed as Pathogenicity Islands (PI). Five such PIs found with sizes ranging from 2kb to 52kb and named PI-1 to PI-5 in the order they were seen on the optical map of Cbm (Figure 2.8). Although no obvious toxicity genes could be found in these regions but further analysis is needed. Also, a 40kb region of the 52kb PI-2, could not be covered by any sequencing data and is the one gap in Cbm genome assembly. On comparing the two genomes using MAUVE, it was seen that Cbp contains 5 of the 8 plasmids in Cbm (Table 2.4), including the large 109kb plasmid. The sequence of 109kb plasmid is similar in both Cbm and Cbp.

Selection of Loss of Activity Cbm mutant colony and bioassays

Cbm was mutagenized using Cisplatin which is a crosslinking agent that forms thymine dimers and induce point and frame shift mutations. Rapid screen of 500 colonies was done using five 3rd instar *Aedes* larvae in 1ml water and loss or decrease of mosquitocidal activity was noted. Nine colonies that showed loss or decrease of toxicity were then grown in 10ml cultures and bioassayed using 25 third-instar larvae in plastic cups containing 200 ml tap water and then fed whole culture of the Cbm-Cisplatin mutant colony at different concentration. Mortality was recorded at 24 h. Only one colony was isolated as mutant colony that showed complete loss of activity. Genomic DNA was extracted and PCR using Cry16A and Cry17A specific primers confirmed the loss of plasmid from the mutant colony.

Presence of Mobile Genetic Elements

Since many of the pathogenic activity in bacteria is based and acquired be plasmids, it was plausible that the larvicidal activity of Cbm was also because of the genes present on plasmids. Genome BGI scaffolds of Cb, Cbp and Cbm were aligned and compared using Mauve™ Progressive Alignment Program (Figure 2.9). It was seen that eight of the Cbm genome scaffolds did not show matching sequences against the non-larvicidal Cb (Figure 2.10). Out of the same eight scaffolds, five scaffolds of Cbp were also absent in Cb genome. This was further validated by PCR data generated in our lab previously by Dr. Qureshi. By combining the PCR data and alignment data, we concluded that there are eight plasmids in Cbm, two of which are also present in Cb

(Table 2.4). Dr. Nadia Qureshi has also previously found a Cbm loss of activity mutant using gamma irradiation and its genomic DNA was isolated and sequenced using Illumina sequencing platform by BGI for us. When the Cbm and Cbm-mutant genome scaffolds were aligned using MAUVE, there were four scaffolds that were absent (Table 2.5) and these were same as those shown to be putative plasmids earlier.

Discussion

Whole genome sequencing is one of the newest strategies available for discovering novel virulence factors and unraveling bacterial pathogenesis (Pallen and Wren, 2007). Although Sanger sequencing has been the traditional method of sequencing, it is quickly being replaced by newer sequencing strategies. These next-generation sequencing (NGS) platforms generate more data in a shorter amount of time and for less cost. They are an attractive option for clostridial genome sequencing and unravel a novel and useful tool for the discovery of completely new insecticidal-toxin genes that would had otherwise proved to be difficult to identify.

Previous work done on isolating the larvicidal activity of Cbm could not identify any factors associated with its toxicity (Juarez-Perez et al., 2001). Since the classical cloning and protein methods did not work, our lab undertook a comparative genomic approach to identify larvicidal toxin genes from Cbm. We sequenced the entire genomes of Cb, Cbp, Cbm and Cbm-mutant. All bacterial strains were sequenced since clostridial genomes tend to vary a lot in size and numbers of genomes and plasmids and also physiology and their environmental conditions. Cbm and Cbp show mosquitocidal

72

activity whereas Cb the type strain is non mosquitocidal. As the genome of any organism reflects its lifestyle, phylogeny and physiology, any differences in regard to the genomic diversity by comparative genomic analysis of these three species could give an indication of the genes or operons or pathogenicity islands that could confer mosquitocidal activity to *C. bifermentans subsp malaysia*. Therefore, the objective was to sequence the entire genomes of Cb, Cbm and Cbp and to compare and analyze them to search for the presence of known toxin genes, as well as for the regions of synteny and divergence.

We used an efficient strategy of combining Illumina, PacBio and 454 de novo sequencing, which allowed the genome to be sequenced completely and assembled correctly with just one gap remaining. Optical maps of Cbm and Cbp were generated to provide a structural backbone for aligning and orienting the scaffolds and the final sequenced genome aligned directly to the restriction map generated by optical mapping. The process used to generate the one-gapped Cbm genome sequence using Cbm optical map and all Cbm sequencing data is shown in Figure 2.5. The NcoI digestion optical map of Cbm genome was first aligned to the in-silico NcoI digestion of BGICbm scaffolds in MapSolver. Of the 39 BGICbm scaffolds, only 9 scaffolds aligned to the map. Next we used a finishing strategy in which BGICbp scaffolds were used to close gaps in Cbm assembly. Briefly, using MAUVE, we aligned the Cbm and Cbp genomes and looked at those single Cbp scaffolds that showed matching sequence to multiple Cbm scaffolds (Figure 2.11). The Cbp Scaffold was therefore used as a "guide" to close the gaps between many Cbm Scaffolds to create one new scaffold. This gap closure was then confirmed by designing PCR primers to the regions that were joined by MAUVE. If PCR

73

showed a correct size product, the gap joining was confirmed. The red arrows in Figure

2.5B indicate where this protocol was followed. The next step was the addition of

BGICbp scaffolds, BGICb scaffolds and Cbm PacBio sequences to the assembly. All the

overlapping sequences were removed keeping those that best aligned the Cbm optical

map. Thus Cbm genome sequence was completed, with one gap remaining.

The comparison of optical maps of Cbm and Cbp also showed a slightly larger

Cbm genome than Cbp due to presence of five regions (PIs) that were present in Cbm and

but absent in Cbp (Figure 2.8). This could be used to explain the difference in the

larvicidal activity of these two bacteria. However, PI-2 which is 52Kb is only partially

covered by all the sequencing data. Hence there is this one gap in the otherwise complete

assembly of Cbm genome.

Earlier reports had not found the presence of a plasmid in Cbm (Selena et al.,

1994). Analysis of the Cb, Cbm and Cbp showed that genomes of larvicidal Cbm and

Cbp differed from Cb by about 250kb (Table 2.4). This could only be explained by

presence of plasmids. Therefore we hypothesized that mosquitocidal activity to Cbm is

associated with a plasmid. To test for this hypothesis we mutagenize Cbm with various

mutagens and select for loss/decrease of larvicidal activity and identify if loss of function

is associated with any plasmid. Dr. Nadia Qureshi in our lab showed the presence of

plasmids by designing PCR primers from the ends of the plasmids and amplifying them

to ascertain which of those scaffolds was a plasmid. She had also generated three loss of

activity Cbm-mutants using gamma irradiation, and genome of one of the mutants was

completely sequenced. I correlated this data with my analysis of the genome scaffolds of

Cb, Cbm, Cbp and Cbm-mutant by MAUVE (Figure 2.9) and found that Cbm contains eight plasmids, two of which are also present in Cb.

The largest plasmid of 109kb was Scaffold 9 in BGI sequencing data (Tables 2.4 and 2.5). Upon analysis of this scaffold annotation given by BGI to us, I found that it contained previously implicated *Cry16A*, *Cry17A*, *Cbm17.1* and *Cbm17.2* genes in an operon. It also has several uncharacterized putative genes, transposons and insertion sequences as well as genes encoding for cell wall associated hydrolases, replication proteins and a type IV secretion system. Apart from all these, there is also an operon encoding for Hemagglutinin (*HA*) gene, Botulinum like toxin gene, Non Toxic Non Hemagglutinin (*NTNH*) gene, *OrfX1*, *OrfX2*, *OrfX3* and *p47* genes. Plasmid map generated by DNAPlotter (Artemis) is given in Figure 2.12. We hypothesize that this plasmid plays an important role in Cbm mosquitocidal activity. This was confirmed when we observed that this plasmid, now named pClosMP was lost in the Cbm-mutant created by Dr. Qureshi. I also mutagenized Cbm with Cisplatin which crosslinks adjacent thymine residues to form thymine dimers to result in mutations. We screened 500 colonies and isolated one loss of larvicidal activity mutant. We confirmed that this mutant had lost pClosMP plasmid by PCR using primers for *Cry16A*, and *Cry17A* genes.

It can be concluded that Cbm is an anaerobic larvicidal bacteria possessing a wide arsenal of virulence factors, mainly on one of its largest plasmids. The annotation shows many know toxin genes present in the circular chromosome and require further study. The genome sequence of *C. bifermentans* subsp. *malaysia* now provides the information on finding completely new toxin genes with novel modes of action to be studied.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M. and Kubal, M. (2008). The RAST server: Rapid annotations using subsystems technology. *BMC Genomics*. **9**:1–15.

Barloy, F., Delecluse, A., Nicolas, L., and Lecadet, M.M. (1996). Cloning and expression of the first anaerobic toxin gene from *Clostridium bifermentans* subsp. *malaysia*, encoding a new mosquitocidal protein with homologies to *Bacillus thuringiensis* delta-endotoxins. *J Bacteriol*. 78: 3099-3105.

Barloy, F., Lecadet, M.M., and Delecluse, A. (1998a). Distribution of clostridial cry-like genes among *Bacillus thuringiensis* and *Clostridium* strains. *Curr Microbiol*. **36**: 232-237.

Barloy, F., Lecadet, M.M., and Delecluse, A. (1998b) Cloning and sequencing of three new putative toxin genes from *Clostridium bifermentans* CH18. *Gene.* **211**: 293-299.

Battisti, L., Green, B. D. & Thorne, C. B. (1985). Mating system for transfer of plasmids among *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*. *J Bacteriol*. **162**: 543-550.

Baum, J. A. & Gonzalez, J. M., Jr. (1992). Mode of replication, size and distribution of naturally occurring plasmids in *Bacillus thuringiensis*. *FEMS Microbiol Lett*. **75**:143-148.

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C. G., Hall, K. P., Evers, D.J., Barnes, C.L., Bignell, H.R., Boutell, J. M., Bryant, J., Carter, R. J., Keira Cheetham, R., Cox, A. J. and Ellis, D. J. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 456 (7218): 53-59.

Ben-Dov, E., Nissan, G., Pelleg, N., Manasherob, R., Boussiba, S. and Zaritsky, A. (1999). Refined, circular restriction map of the *Bacillus thuringiensis* subsp. *israelensis* plasmid carrying the mosquito larvicidal genes. *Plasmid*. **42**(3):186-91.

Berry, C., O'Neil, S., Ben-Dov, E., Jones, A. F., Murphy, L., Quail, M. A., Holden, M. T. G., Harris, D., Zaritsky, A. and Parkhil, J. (2002). Complete Sequence and Organization of pBtoxis, the Toxin-Coding Plasmid of *Bacillus thuringiensis* subsp. *israelensis*. *Appl Environ Microbiol*. **68**(10): 5082–5095.

Carlson, C. R., Caugant, D. A. and Kolsto, A. B. (1994). Genotypic Diversity among *Bacillus cereus* and *Bacillus thuringiensis* Strains. *AEM*. **60**:1719-1725.

Carlson, C. R., Johansen, T. and Kolsto, A. B. (1996). The chromosome map of *Bacillus thuringiensis* subsp. *canadensis* HD224 is highly similar to that of the *Bacillus cereus* type strain ATCC 14579. *FEMS Microbiol Lett.* **141**: 163-167.

Carlson, C. R. and Kolsto, A. B. (1993). A complete physical map of a *Bacillus thuringiensis* chromosome. *J Bacteriol*. **175**: 1053-1060.

Crickmore, N., Zeigler, D.R., Schnepf, E., van Rie, J., Lereclus, D., Baum, J., Bravo, A. and Dean, D.H. (2014). Bacillus thuringiensis Toxin Nomenclature. Available online: http://www.lifesci. sussex.ac.uk/Home/Neil_Crickmore/Bt/

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F. and Cicero, R. (2009). Real-time DNA sequencing from single polymerase molecules. *Science*. **323**:133-138.

Faust, R. M., Abe, K., Held, G. A., Iizuka, T., Bulla, L. A. and Meyers, C. L. (1983). Evidence for plasmid-associated crystal toxin production in *Bacillus thuringiensis* subsp. *israelensis*. *Plasmid*. **9**(1):98-103.

Fonknechten, N., Chaussonnerie, S., Tricot, S., Lajus, A., Andreesen, J. R., Perchat, N., Pelletier, E., Gouyvenoux, M., Barbe, V., Salanoubat, M., Le Paslier, D., Weissenbach, J., Cohen, G. N. and Kreimeyer, A. (2010). *Clostridium sticklandii*, a specialist in amino acid degradation:revisiting its metabolism through its genome sequence. *BMC Genomics*. **11**: 555.

Ferré, J. and van Rie, J. (2002). Biochemistry and genetics of insect resistance to *Bacillus thuringiensis*. *Annu. Rev. Entomol*. **47**: 501–533.

Gonzalez, J. M. and Carlton, B. C. (1980). Patterns of plasmid DNA in crystalliferous and acrystalliferous strains of *Bacillus thuringiensis. Plasmid*. **3**: 92-98.

Green, B. D., Battisti, L. and Thorne, C. B. (1989). Involvement of Tn4430 in transfer of *Bacillus anthracis* plasmids mediated by *Bacillus thuringiensis* plasmid pXO12. *J. Bacteriol*. **171**: 104-113.

Juarez-Perez, V., and Delecluse, A. (2001). The Cry toxins and the putative hemolysins of *Clostridium bifermentans* serovar *malaysia* are not involved in mosquitocidal activity. *J Invertebr Pathol*. **78**: 57-58

Korlach, J., Marks, P.J., Cicero, R.L., Gray, J.J., Murphy, D.L., Roitman, D.B., Pham, T.T., Otto, G.A., Foquet, M. and Turner, S.W. (2008). Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures. *PNAS USA*. **105**(4):1176-1181

Kronstad, J. W., Schnepf, H. E. and Whiteley, H. R. (1983). Diversity of locations for *Bacillus thuringiensis* crystal protein genes. *J Bacteriol*. 154: 419-428.

Lereclus, D., Lecadet, M. M., Ribier, J. & Dedonder, R. (1982). Molecular relationships among plasmids of *Bacillus thuringiensis*: conserved sequences through 11 crystalliferous strains. *Mol Gen Genet*. **186**: 391-398.

Margulies, M., Egholm, M., Altman, W. E., Attiya,S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S. and Ho, C. H. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 437(7057):376-380.

Miyamoto, K., D. J. Fisher, D. J., Li, J., Sayeed, S., Akimoto, S. and McClane, B. A. (2006). Complete sequencing and diversity analysis of the enterotoxin-encoding plasmids in *Clostridium perfringens* type A non-food-borne human gastrointestinal disease Isolates. *J Bacteriol*. **188**:1585-1598.

Myers, G. S., Rasko, D. A., Cheung, J. K., Ravel, J., Seshadri, R., DeBoy, R. T. , Ren, Q., Varga, J., Awad, M. M., Brinkac, L. M., Daugherty, S. C., Haft, D. H., Dodson, R. J., Madupu, R., Nelson, W. C., Rosovitz, M. J., Sullivan, S. A., Khouri, H., Dimitrov, G. I., Watkins, K. L., Mulligan, S., Benton, J., Radune, D., Fisher, D. J., Atkins, H. S., Hiscox, T., Jost, B. H., Billington, S.J., Songer, J. G., McClane, B. A., Titball, R. W., Rood, J. I., Melville, S. B. and Paulsen, I. T. (2006). Skewed genomic variability in strains of the toxigenic bacterial pathogen, *Clostridium perfringens*. *Genome Res*. **16**:1031-1040.

Oppert, B.; Kramer, K.J.; Beeman, R.W.; Johnson, D.; McGaughey, W.H. Proteinase-mediated insect resistance to *Bacillus thuringiensis* toxins. *J. Biol. Chem*. **272**: 23473–23476.

Ravin N. V., Kuprianov V. V., Gilcrease E. B., Casjens S. R. (2003). Bidirectional replication from an internal ori site of the linear N15 plasmid prophage. *Nucleic Acids Res*. **31**:6552–6560

Rood, J. I. (1998). Virulence genes of *Clostridium perfringens*. *Annu Rev Microbiol*. **52**:333-360.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *PNAS USA*. **74**:5463-5467.

Sayyed, A.H., Haward, R., Herrero, S., Ferré, J. and Wright, D.J. (2000). Genetic and biochemical approach for characterization of resistance to *Bacillus thuringiensis* toxin Cry1Ac in a field population of the diamondback moth, *Plutella xylostella*. *Appl. Environ. Microbiol*. **66**: 1509–1516.

Seleena, P. and Lee, H. P. (1994). Absence of plasmid in mosquitocidal *Clostridium bifermentans* serovar *malaysia*. *S.E. Asian J. Trop. Med. Pub. Health* **25**(2): 394-396

Shendure, J. and Ji, H. (2008). Next-generation DNA sequencing. *Nat. Biotechnol*. **26**: 1135–1145.

Shimizu, T., K. Ohtani, H. Hirakawa, K. Ohshima, A. Yamashita, T. Shiba, N. Ogasawara, M. Hattori, S. Kuhara, and H. Hayashi. (2002). Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater. *PNAS*. **99**:996-1001.

Thie´ry, I., S. Hamon, B. Gaven, and H. de BarjaC. (1992). Host range of *Clostridium bifermentans* serovar *malaysia*, a mosquitocidal anaerobic bacterium. *J. Am. Mosq. Control Assoc*. **8**:272–277

Turcatti, G., Romieu, A., Fedurco, M. and Tairi, A.P. (2008). A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Nucleic Acids Research*. **36**(4):e25.

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

Ward, E. S. and Ellar, D. J. (1983). Assignment of the δ-endotoxin gene of *Bacillus thuringiensis* var. *israelensis* to a specific plasmid by curing analysis. *FEBS Letters*. **158**: 45-49.

Young, M., Staudenbauer, W. L., and Minton, N. P. (1989). Genetics of *Clostridium*. In "Clostridia" (N. P. Minton and D. J. Clarke, Eds.), pp. 63–104. Plenum, New York.

Zhou, S., Herscheleb, J. and Schwartz, D.C. (2007). A Single Molecule System for Whole Genome Analysis. New high throughput technologies for DNA sequencing and genomics. *Elsevier*. pp. 269–304.

Species	Group	Subtype	Strain	Sequence	Size (kb)	Accession
C. botulinum	I	A1	ATCC 3502	chr[a]	3,887	AM412317
	I			plasmid	16	AM412317
C. botulinum	I	A1	ATCC 19397	chr[a]	3,863	CP000726
C. botulinum	I	A1	Hall	chr[a]	3,761	CP000727
C. botulinum	I	A1(B)	NCTC 2916	chr[a]	4,031	ABDO02000001-49
C. botulinum	I	A2	Kyoto-F	chr[a]	4,155	CP001581
C. botulinum	I	A3	Loch Maree	chr	3,993	CP000962
	I			plasmid[a]	267	CP000963
C. botulinum	I	B5a4	strain 657	chr	3,978	CP001083
C. botulinum				plasmid[a]	270	CP001081
C. botulinum	I			plasmid	10	CP001082
C. botulinum	I	A5	H04402 065	chr[a]	3,920	FR773526
C. botulinum	I	B1	Okra	chr	3,958	CP000939
	I			plasmid[a]	149	CP000940
C. botulinum	I	B5f2	—	plasmid	4,200	ABDP01000001-70
C. botulinum	I	F1	Langeland	chr[a]	3,995	CP000728
C. botulinum	I			plasmid	18	CP000729
C. botulinum	I	F1	F230613	chr[a]	3,993	CP002011
C. botulinum	I			plasmid	18	CP002012
C. botulinum	II	B4	Eklund 17B	chr	3,800	CP001056
C. botulinum	II			plasmid[a]	48	CP001057
C. botulinum	II	E1	Beluga	chr[a]	4,000	ACSC01000001-6
C. botulinum	II	E3	E43	chr[a]	3,660	CP001078
C. botulinum	III	C	Stockholm	chr	2,660	AESA00000000
				phage[a]	55	AP008983
C. botulinum	III	C	Eklund	chr, phage[a]	2,961	ABDQ01000001-76
C. botulinum	III	CD	08-BKT015925	chr, phage[a]	2,773	CP002410
				plasmid	203	CP002411
C. botulinum	III	CD	08-V891	chr	3,139	AESC00000000
C. botulinum	III	D	1873	chr, phage[a]	2,379	ACSJ01000001-19
				plasmid	108	CP001659
				plasmid	54	CP001660
C. butyricum	VI	E4	BL5262	chr[a]	4,758	ACOM01000001-13
C. butyricum	VI	E4	BL5521	chr[a]	4,541	ABDT01000001-123

Figure 2.1: A list of completely sequenced *Clostridial* genomes. The Group designation, strain information and chromosome (chr), plasmid or phage sequence related to the GenBank accession is provided. Picture taken from Rummel and Binz, 2013.

[a] bont gene cluster locations
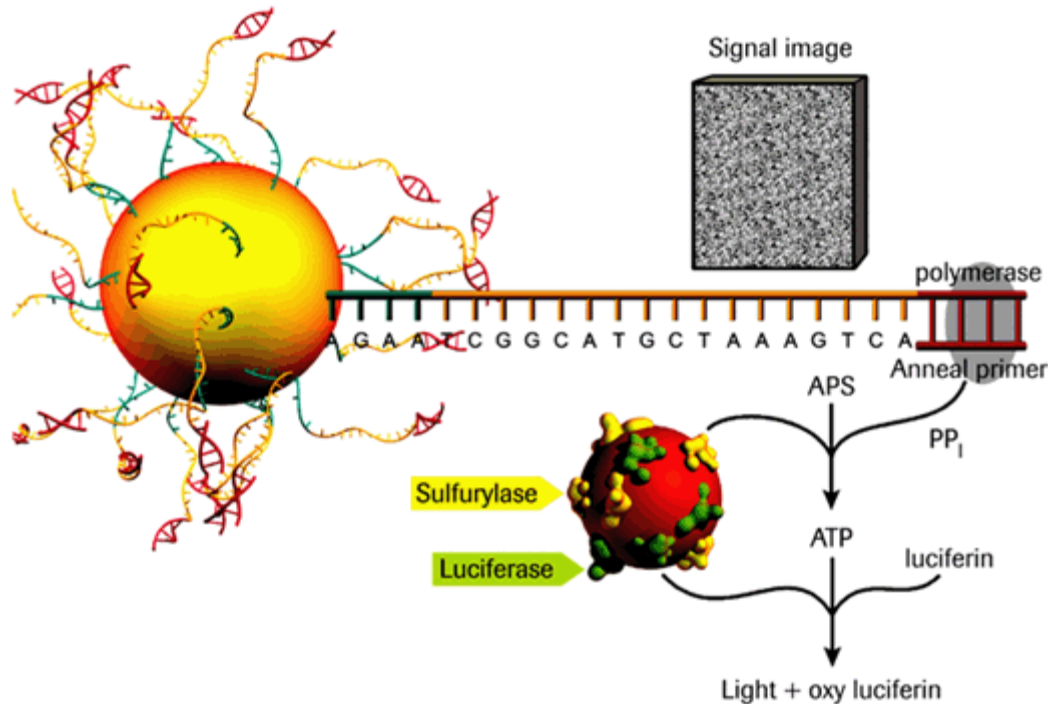
Figure 2.2: The pyrosequencing process. One of four nucleotides is washed sequentially over copies of the sequence to be determined, causing polymerases to incorporate complementary nucleotides. The incorporation stops if the longest possible stretch of the available nucleotide has been synthesized. In the process of incorporation, one pyrophosphate per nucleotide is released and converted to ATP by an ATP sulfurylase. The ATP drives the light reaction of luciferases present and a light signal proportional (within limits) to the number of nucleotide incorporations can be measured. Taken from 454Roche company website www.454.com
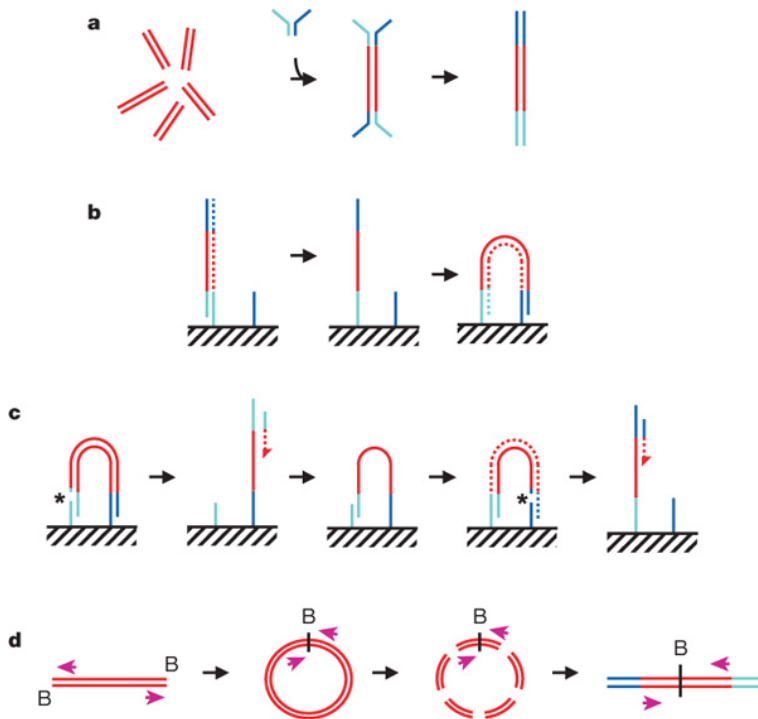
Figure 2.3 Reversible terminator chemistry applied by the Illumina sequencing.
Sequencing primers are annealed to the adapters of the sequences to be determined.
Polymerases are used to extend the sequencing primers by incorporation of fluorescently
labeled and terminated nucleotides. The incorporation stops immediately after the first
nucleotide due to the terminators. The polymerases and free nucleotides are washed away
and the label of the bases incorporated for each sequence is read with four images taken
through different filters and using two different lasers (red: A, C and green: G, T) to
illuminate fluorophores. Subsequently the fluorophores and terminators are removed and
the sequencing continued with the incorporation of the next base. Taken from Bentley et
al., 2008.
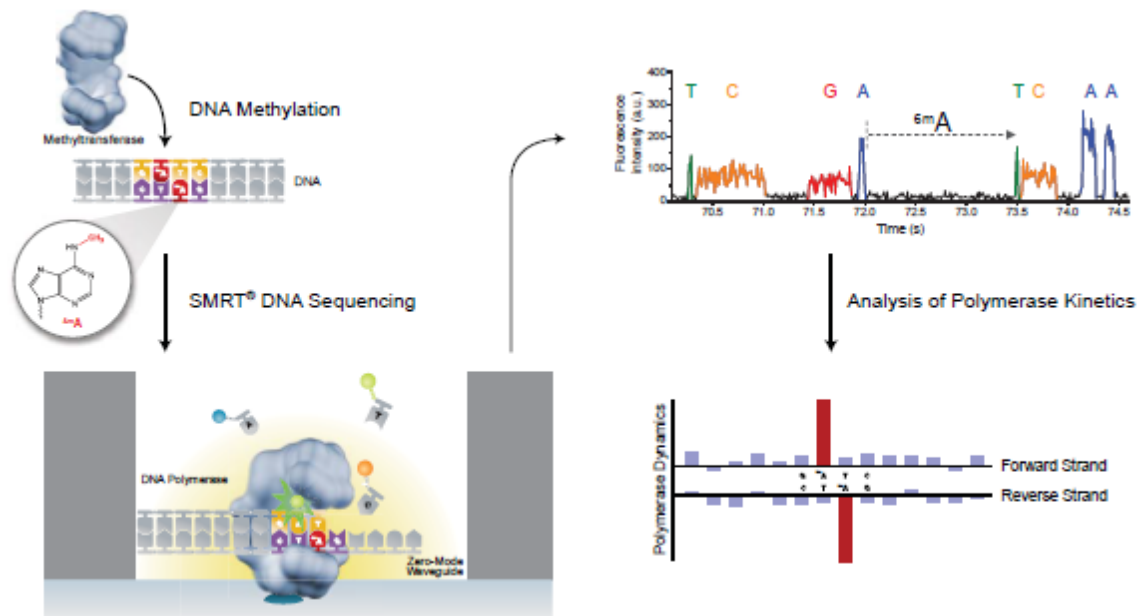
Figure 2.4: Pacbio® RS II system. Methyltransferases bind specifically to DNA motifs in a genome and methylate bases. Pacbio software locates modified sites and motifs. Their SMRT technology enables the observation of DNA synthesis as it occurs in real time. Taken from www.pacificbiosciences.com
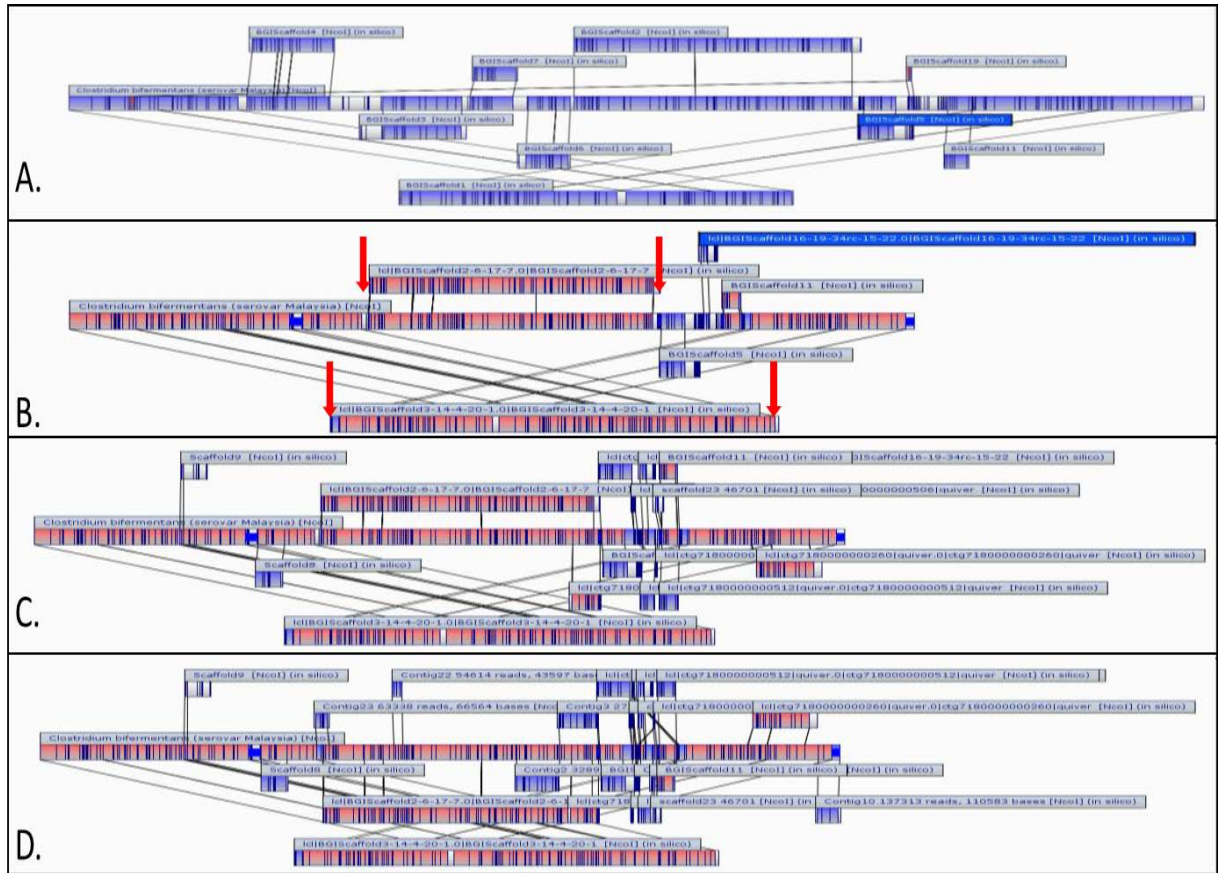
Fig 2.5: Alignments between the whole-genome optical maps and the in silico genome sequence assemblies at various stages of the project. Dark blue represents cut sites, light blue regions indicate alignment, orange regions indicate comparative sequences and white regions indicate no alignment.

(A) Early comparison of an optical map derived from NcoI digestion of the Cbm genome to the assembled BGICbm scaffolds generated by illumina sequencing. Of the 39 scaffolds generated by sequencing only 9 aligned to the map.

(B) The Cbm optical map derived from NcoI digestion of the chromosome is presented as a single contig in the center. The finishing strategy including closing gaps using MAUVE and confirming the gap closure by PCR. Red arrows indicate where this protocol was followed.

(C) Addition of BGICbp scaffolds, BGICb scaffolds and Cbm PacBio sequences were added to the assembly.

(D) All the overlapping sequences were removed keeping those that best aligned the Cbm optical map.

84

KEGG pathway classification
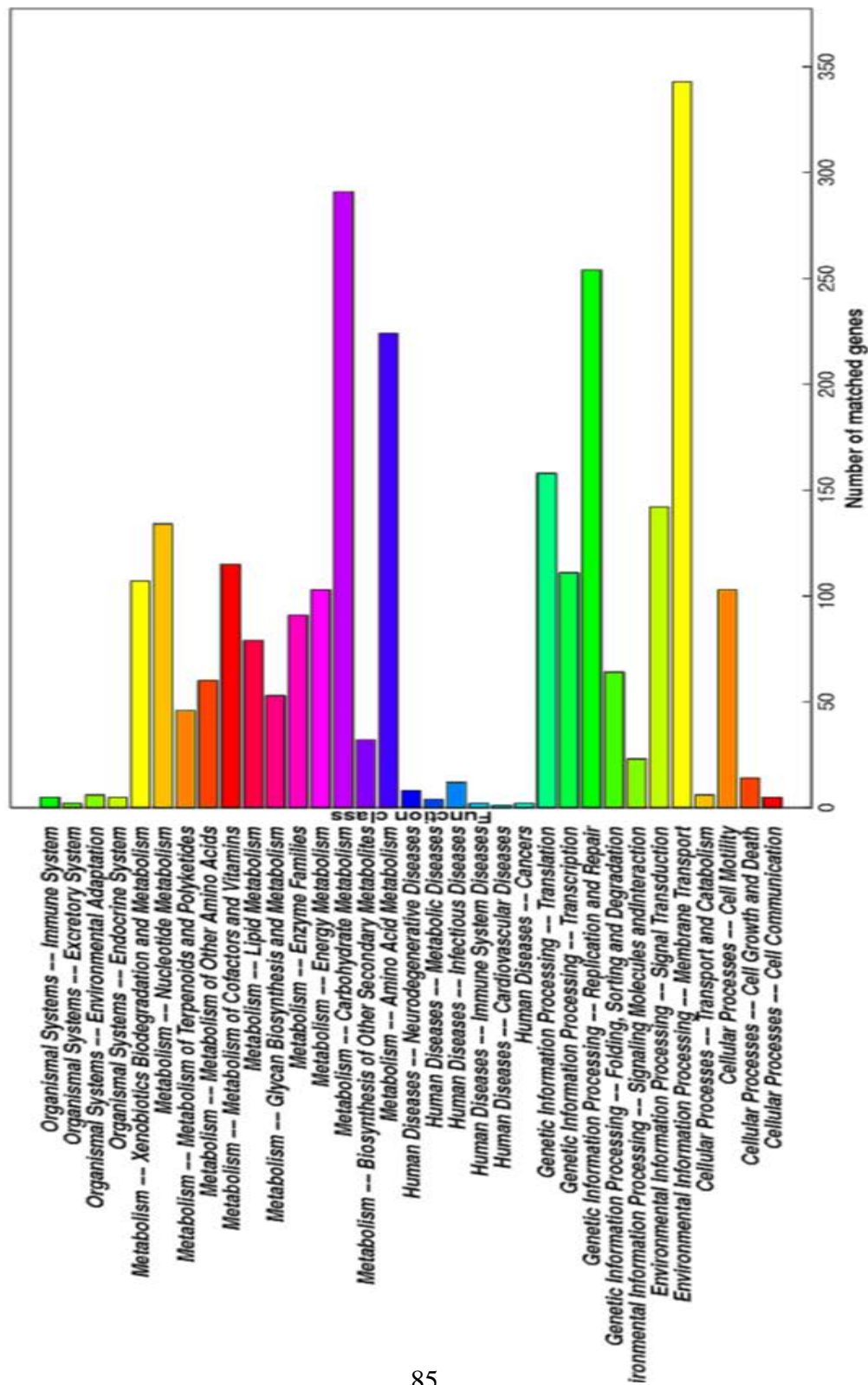
Figure 2.6: KEGG pathway classification of Cbm. Data generated by BGI, China. KEGG provides molecularinteraction and metabolic pathway. The histogram shows the distribution of functions of putative genes in Cbm genome.

KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from genomic and molecular-level information. It is a computer representation of the biological system, consisting of molecular building blocks of genes and proteins (genomic information) and chemical substances (chemical information) that are integrated with experimental knowledge on such systemic functions from literature. These databases constitute the reference knowledge base for biological interpretation of genomes and high-throughput molecular datasets.

COG function classification

B: Chromatin structure and dynamics
C: Energy production and conversion
D: Cell cycle control, cell division, chromosome partitioning
E: Amino acid transport and metabolism
F: Nucleotide transport and metabolism
G: Carbohydrate transport and metabolism
H: Coenzyme transport and metabolism
I: Lipid transport and metabolism
J: Translation, ribosomal structure and biogenesis
K: Transcription
L: Replication, recombination and repair
M: Cell wall/membrane/envelope biogenesis
N: Cell motility
O: Posttranslational modification, protein turnover, chaperones
P: Inorganic ion transport and metabolism
Q: Secondary metabolites biosynthesis, transport and catabolism
R: General function prediction only
S: Function unknown
T: Signal transduction mechanisms
U: Intracellular trafficking, secretion, and vesicular transport
V: Defense mechanisms

Figure 2.7: COG function classification of Cbm. Data generated by BGI, China.
The histogram represents the functional predictions of the putative genes in Cbm
genome.

The Clusters of Orthologous Groups of proteins (COGs) database has been designed as
an attempt to classify proteins from completely sequenced genomes on the basis of the
orthology concept. Orthologs are direct evolutionary counterparts related by vertical
descent as opposed to paralogs which are genes within the same genome related by
duplication. Typically, orthologous proteins have the same domain architecture and the
same function although there are significant exceptions and complications to this
generalization.
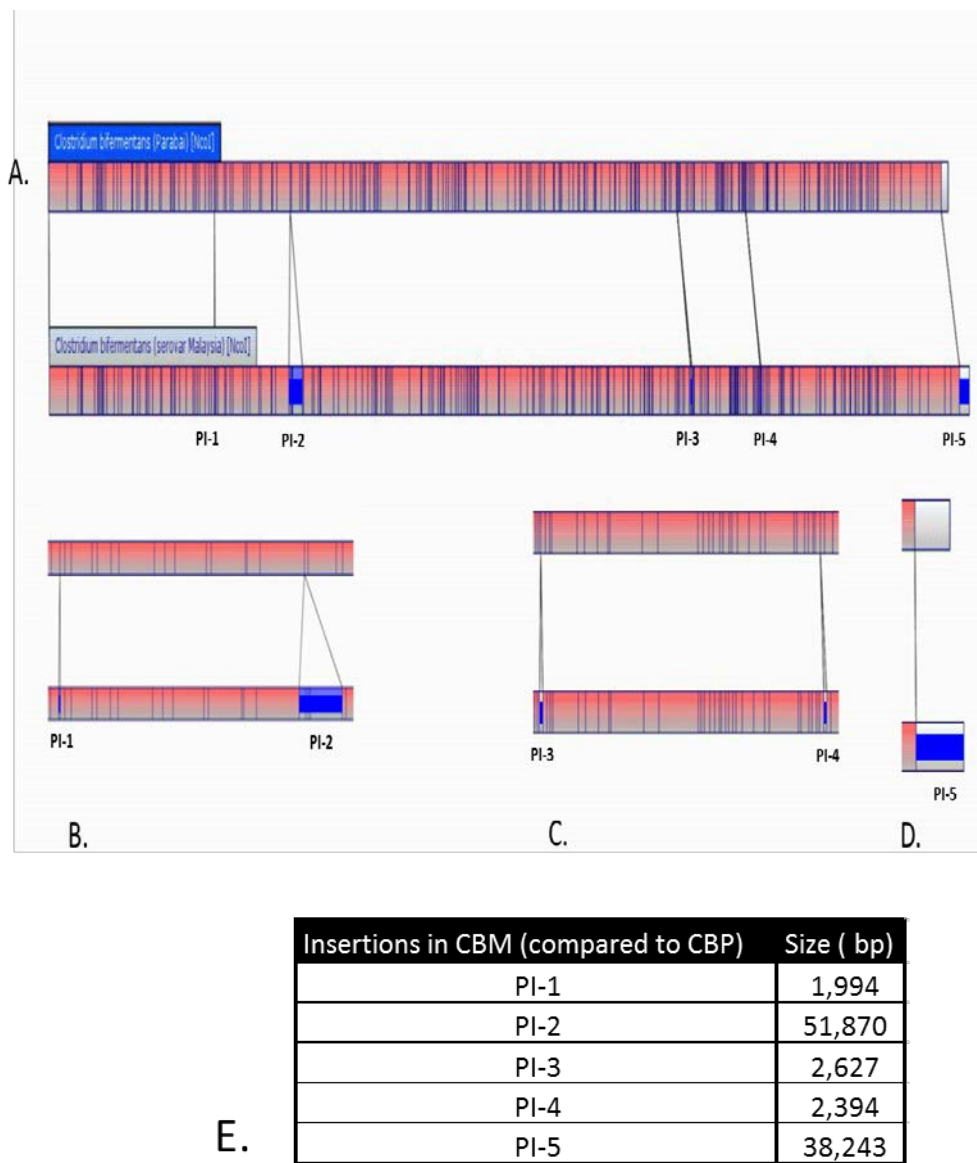
Figure 2.8: Comparison of chromosome organization of Cbm and Cbp showing Pathogenecity Islands in Cbm.

    A. depicts the difference in sizes of Cbm and Cbp genomic optical maps is due to insertion of five regions termed PI for putative Pathogenecity Islands
    B. depicts the expanded view of PI-1 and PI-2
    C. depicts the expanded view of PI-3 and PI-4
    D. depicts the expanded view of PI-5
    E. Tabular depiction of sizes of the five PI's

Cb

Cbm

Cbp
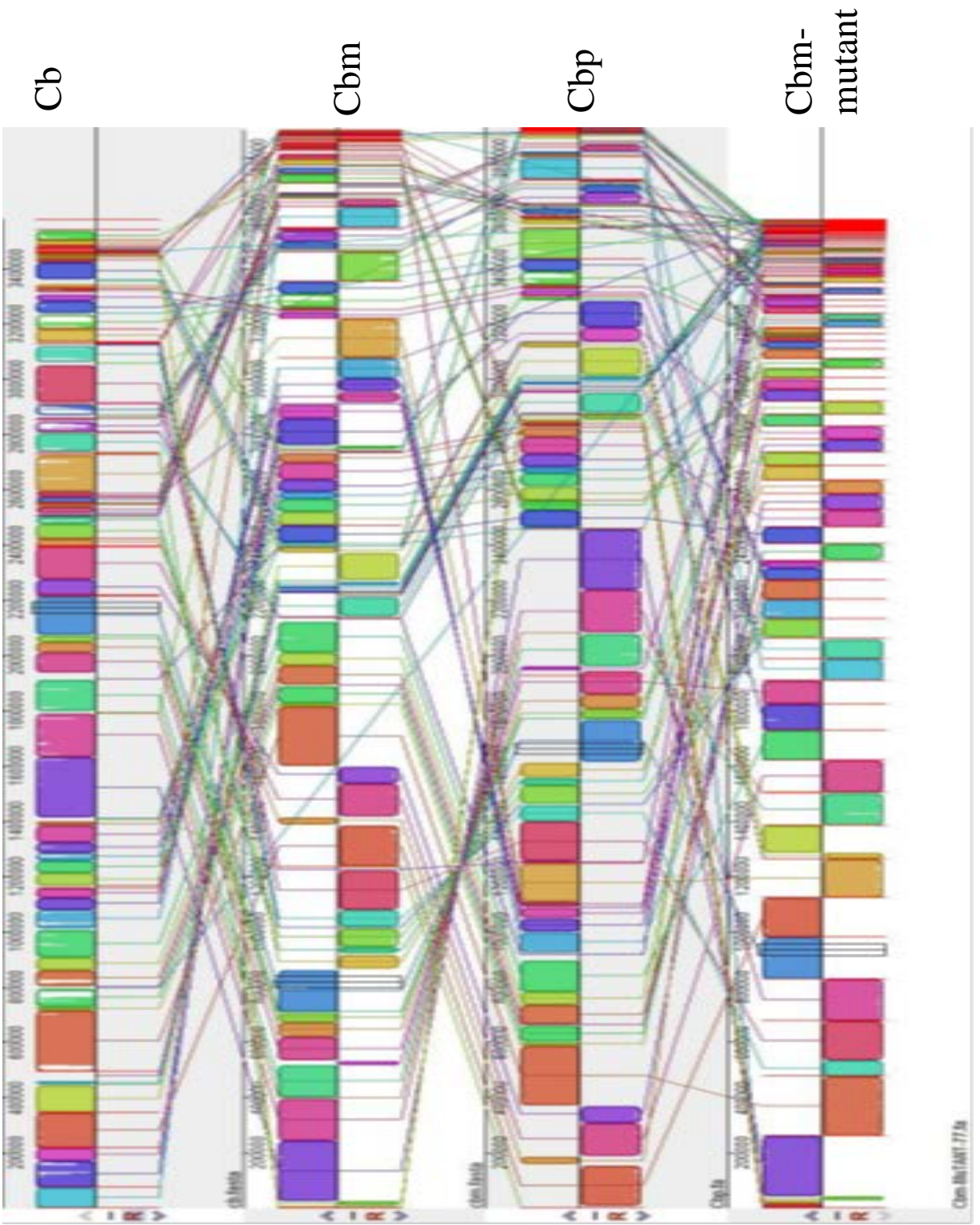
Cbm-mutant

Figure 2.9. A whole-genome alignment of Cb, Cbm, Cbp and Cbm-mutant genomes.
MAUVE algorithm (Darling et al., 2010) was used for the alignment of the four genome
sequences of Cb, Cbm, Cbp and Cbm-mutant. BGI sequencing data was used. This is a
screenshot of Mauve visualization of same sequence blocks identified between the
chromosomes of Cb, Cm, Cbp and Cbm-mutant. Each of the same sequence blocks has a
same color. Red vertical bars demarcate scaffold boundaries. The Mauve rearrangement
viewer enables users to interactively zoom in on regions of interest and examine the local
rearrangement structure. The genomes were drawn to scale based on the reference Cb
genome (The first genome is always used as the reference genome).
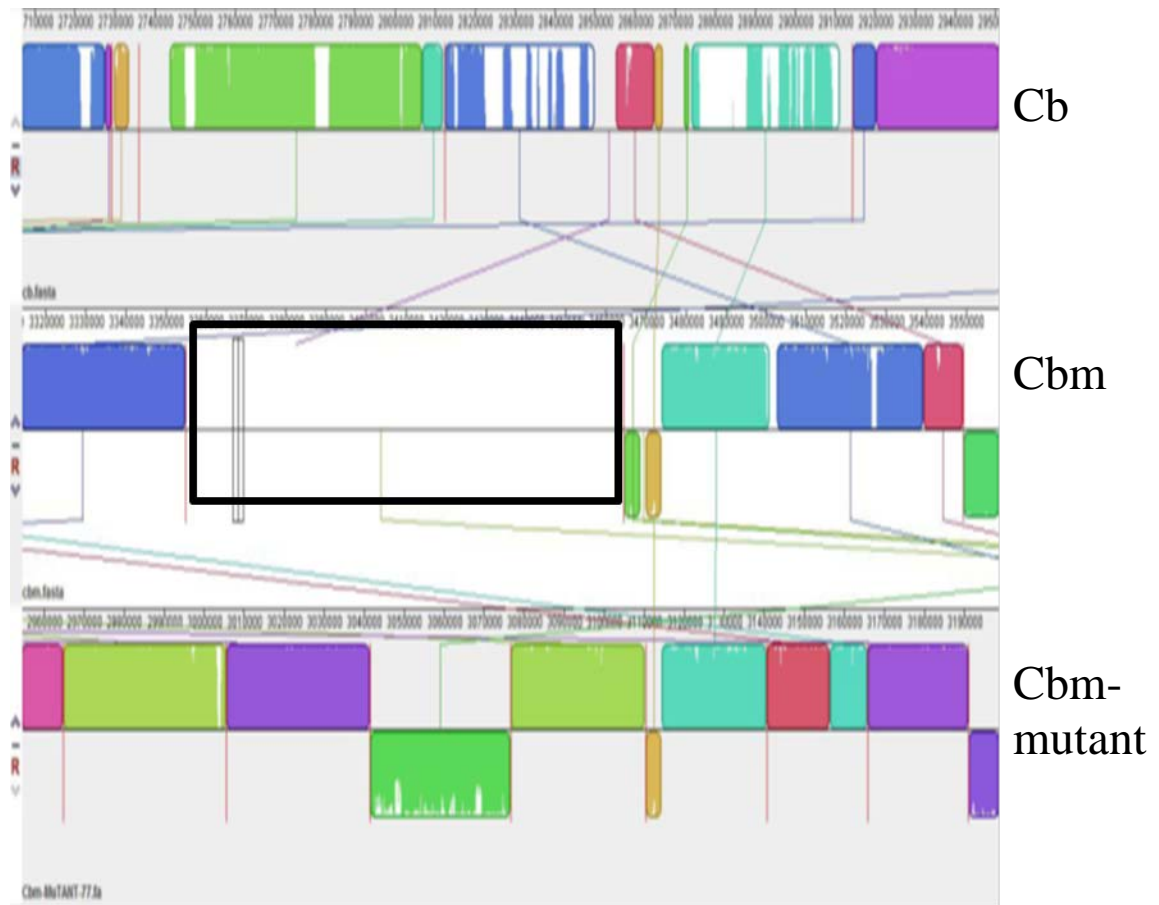
Figure 2.10. Comparative analysis of Cb, Cbm and Cbm-mutant genomes. The BGI consensus sequence of each strain was used for comparison by MAUVE version 2.3.1. This is a close-up depiction of a 109kb plasmid (Scaffold 9 in BGICbm sequence) in comparison to non-toxic Cb and Cbm-mutant reference genomes. The white areas indicated low or zero reads. In this example, we can see there are no matching sequences to Cbm Scaffold 9 (indicated by rectangle) in either Cb or Cbm-mutant genomes.
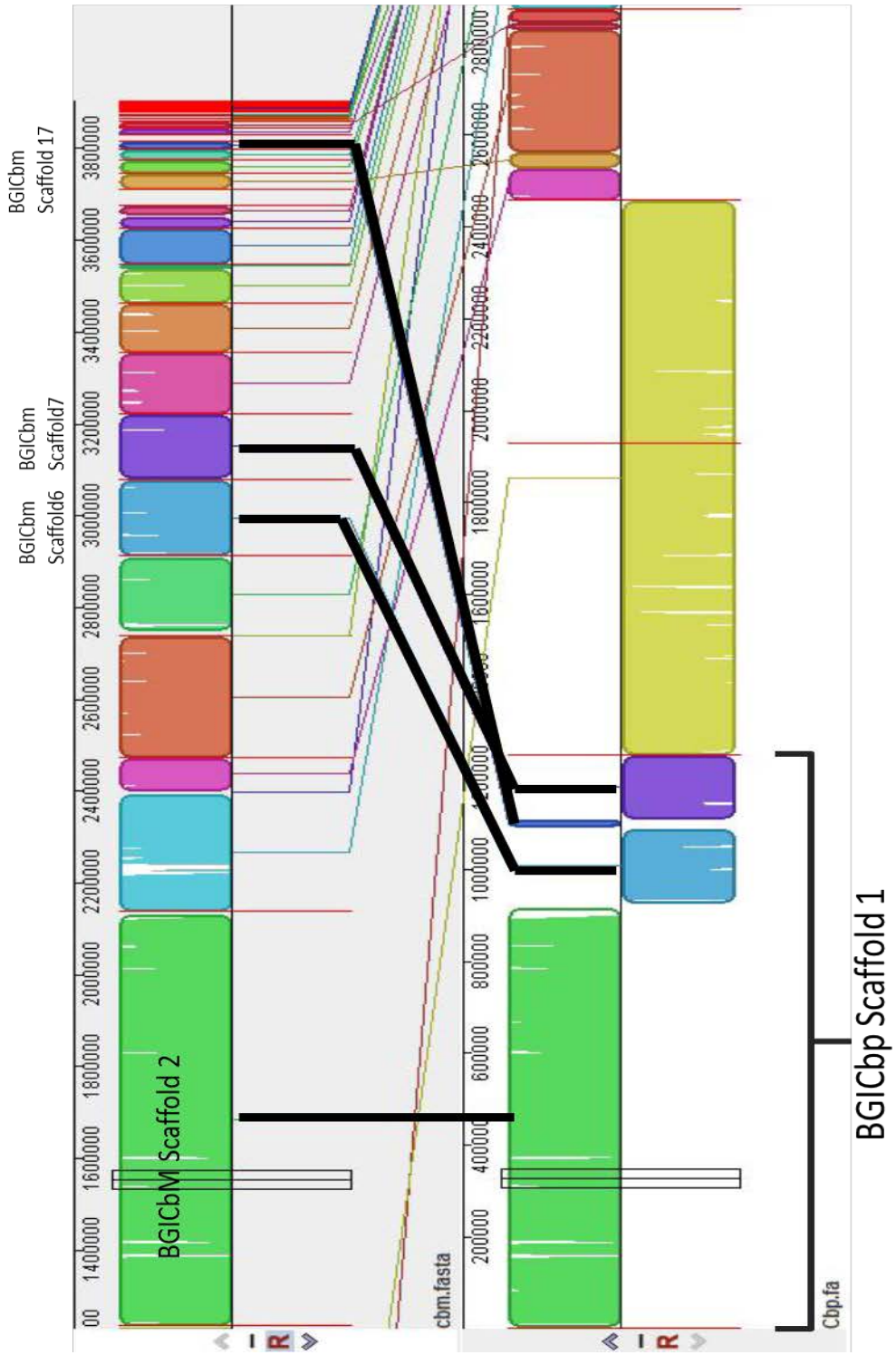
Figure 2.11. Comparative analysis of Cbm and Cbp genomes. The BGI consensus sequence of each strain was used for comparison by MAUVE version 2.3.1. This is a close-up depiction of a BGICbp Scaffold 1being used to close gaps in Cbm assembly. The white areas indicated low or zero reads. In this example, we can see Cbp Scaffold 1 sequence shows the orientation of assembly Cbm Scaffolds 2, 6, 7 and 17. Cbp Scaffold 1 was therefore used as a "guide" to close the gaps between Cbm Scaffolds 2, 6, 7 and 17 to join them and create one new single Scaffold 2-6-17-7. This was the finishing strategy of closing gaps using MAUVE and confirming the gap closure by PCR as depicted in Figure 2.5B.
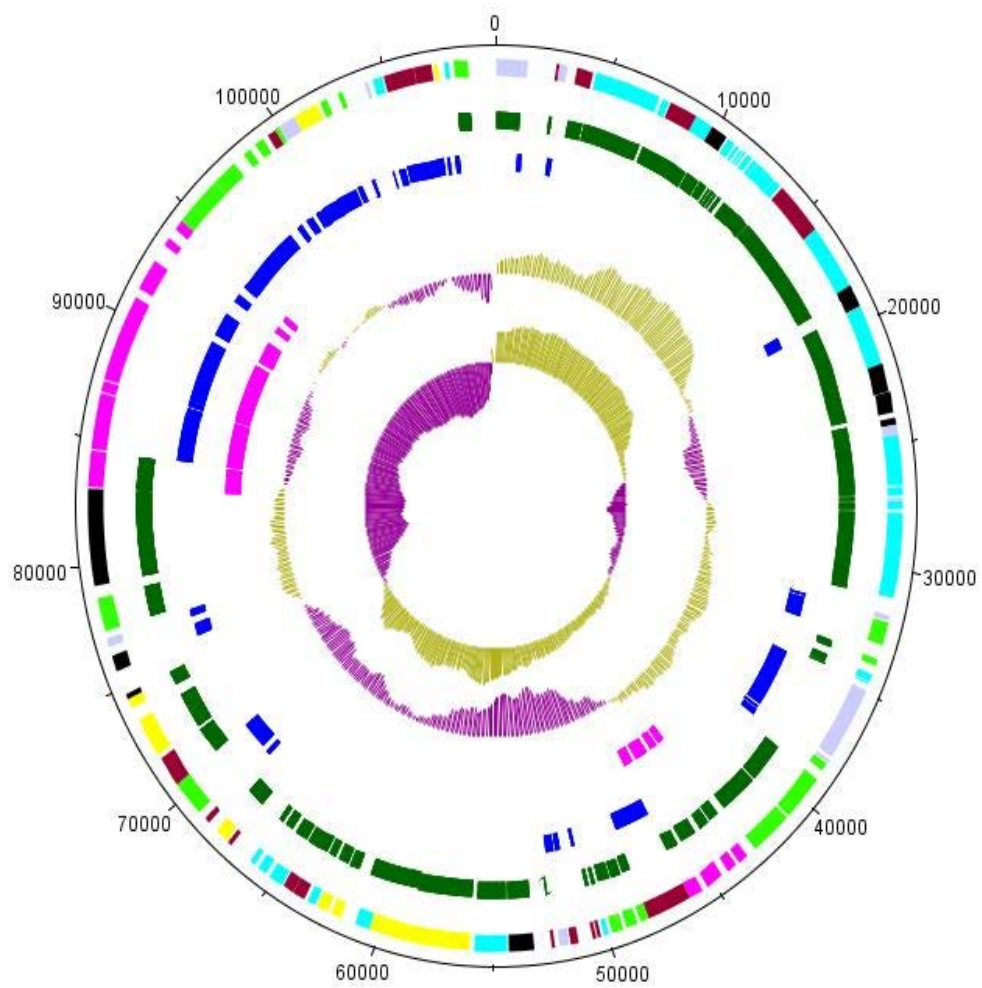
Figure 2.12: Circular representation of 109 Kb plasmid pClosMP.
The inner circle   represents GC bias [(G − C)/(G + C)], with positive values in Beige
                    and negative values in purple;
The second circle represents G+C content;
The third circle   represents the toxin containing operons (pink);
 The fourth circle represent predicted genes on the reverse strand(dark blue);
 The fifth circle   represent predicted genes on the forward strand(dark green)
 The sixth circle   represents the complete plasmid predicted genes (color coded)

Color coding for the genes is as follows:
gray =   regulatory;
pink =   toxin operons;
blue =   conserved hypothetical;
red =   unknown;
green = transposon related ; mobile elements;
Black = cell wall associated; surface associated;
yellow = miscellaneous metabolic genes.

The outer scale is marked in bases.

The origin of replication of plasmid is set at 0. It coincides with the GC ratio change.

Table 2.1: Annotation of genes encoded by the pClosMP

Gene No.	Gene Name	Position on Plasmid (in bp)		Length (in bp)	For/ Rev	Putative Function
		Start	End			
1	pClosMP_1	7	1173	1166	+	initiator RepB protein [*Clostridium perfringens*]
2	pClosMP_2	1207	1419	212	-	Cro/CI family transcriptional regulator - like protein [*Clostridium botulinum*]
3	pClosMP_3	1745	1924	179	+	NONE
4	pClosMP_4	2014	2454	440	-	transcriptional regulator [*Clostridium kluyveri*]
5	pClosMP_5	2600	2749	149	+	NONE
6	pClosMP_6	2802	3092	290	+	hypothetical protein pxo2_34 [*Bacillus anthracis*]
7	pClosMP_7	3561	4265	704	+	putative lipoprotein [*Clostridium bifermentans*]
8	pClosMP_8	4357	6708	2351	+	hypothetical protein pls20_p011 [*Bacillus subtilis subsp. natto*]
9	pClosMP_9	6749	7246	497	+	NONE
10	pClosMP_10	7376	7744	368	+	hypothetical protein pls20_p014 [*B. subtilis subsp. natto*]
11	pClosMP_11	7769	8986	1217	+	type II/IV secretion system protein [*Bacillus subtilis subsp. natto*]
12	pClosMP_12	9000	9812	812	+	hypothetical protein pls20_p016 [*Bacillus subtilis subsp. natto*]
13	pClosMP_13	9815	10549	734	+	hypothetical protein pls20_p017 [*Bacillus subtilis subsp. natto*]
14	pClosMP_14	10590	10952	362	+	hypothetical protein pls20_p018 [Bacillus subtilis subsp. natto]
15	pClosMP_15	11054	11197	143	+	conserved hypothetical protein [*Bacillus cereus*]

16	pClosMP_16	11269	11409	140	+	hypothetical protein Tter_1152 [*Thermobaculum terrenum*]
17	pClosMP_17	11447	11590	143	+	hypothetical protein [*Thermobaculum terrenum*]
18	pClosMP_18	11682	11906	224	+	hypothetical protein pls20_p023 [*Bacillus subtilis subsp. natto*]
19	pClosMP_19	12034	12159	125	+	hypothetical protein pls20_p024 [*Bacillus subtilis subsp. natto*]
20	pClosMP_20	12180	12632	452	+	hypothetical protein pls20_p025 [*Bacillus subtilis subsp. natto*]
21	pClosMP_21	12622	13191	569	+	NONE
22	pClosMP_22	13196	13543	347	+	hypothetical protein PSHAa2415 [*Pseudoalteromonas haloplanktis TAC125*]
23	pClosMP_23	13602	15764	2162	+	virD4 type IV secretion system protein VirD [*B. subtilis subsp. natto*]
24	pClosMP_24	15787	18480	2693	+	Type V secretory pathway, adhesin AidA [*Bacillus subtilis subsp. subtilis*]
25	pClosMP_25	18464	18787	323	+	hypothetical protein pls20_p028 [*Bacillus subtilis subsp. natto*]
26	pClosMP_26	18777	19355	578	+	hypothetical protein pls20_p029 [*Bacillus subtilis subsp. natto*]
27	pClosMP_27	19370	21265	1895	+	putative ATPase TraE [*B. subtilis subsp. natto*]
28	pClosMP_28	21267	21839	572	+	VirB4 component of type IV secretory pathway [*B. subtilis subsp. natto*]
29	pClosMP_29	21832	22926	1094	+	probable cell wall-binding protein [*B. subtilis natto*]
30	pClosMP_30	22962	23765	803	+	replication-associated protein [*Clostridium difficile ATCC 43255]*

31	pClosMP_31	24011	24259	248	+	hypothetical protein pls20_p034 [*Bacillus subtilis subsp. natto*]
32	pClosMP_32	24264	24653	389	+	hypothetical protein pls20_p036 [*Bacillus subtilis subsp. natto*]
33	pClosMP_33	24682	26676	1994	+	ATPase involved in DNA repair [*Bacillus subtilis subsp. natto*]
34	pClosMP_34	26719	27048	329	+	hypothetical protein pls20_p038 [Bacillus subtilis subsp. natto]
35	pClosMP_35	27118	27291	173	+	hypothetical protein CV_0334 [*Chromobacterium violaceum* ATCC 12472]
36	pClosMP_36	27309	27626	317	+	hypothetical protein CdifA_16466 [*Clostridium difficile* ATCC 43255]
37	pClosMP_37	27754	30537	2783	-	single-stranded-DNA-specific exonuclease RecJ [*Listeria ivanovii*]
38	pClosMP_38	30527	31147	620	-	transposase [*Clostridium perfringens* D str. JGS1]
39	pClosMP_39	31855	32004	149	+	transposon Tn21 resolvase [*Clostridium perfringens*B]
40	pClosMP_40	32077	33027	950	+	conserved hypothetical protein [*Culex quinquefasciatus*]
41	pClosMP_41	33594	33908	314	+	hypothetical protein AC3_A0058 [*Clostridium perfringens* E str]
42	pClosMP_42	34283	34585	302	-	DNA-binding protein [*Clostridium difficile* 630]
43	pClosMP_43	34631	34744	113	-	DNA polymerase III, beta subunit [*Clostridium perfringens* D str.]
44	pClosMP_44	34944	37910	2966	-	transposase for transposon Tn1546 [*Enterococcus faecium*]
45	pClosMP_45	38045	38212	167	+	transposon Tn1546 resolvase [*Enterococcus faecium*]

46	pClosMP_46	38264	38614	350	+	transposon Tn1546 resolvase [*Enterococcus faecium*]
47	pClosMP_47	39014	40867	1853	+	pesticidal crystal-like protein cry16Aa [*Clostridium bifermentans malaysia*]
48	pClosMP_48	40952	42808	1856	+	pesticidal crystal-like protein cry17Aa [*Clostridium bifermentans malaysia*]
49	pClosMP_49	43234	43695	461	+	hemolysin-like protein [*Clostridium bifermentans malaysia*]
50	pClosMP_50	43829	44287	458	+	hemolysin-like protein [*Clostridium bifermentans malaysia*]
51	pClosMP_51	44569	45369	800	-	NONE
52	pClosMP_52	45607	46185	578	+	resolvase homolog yefC [*Bacillus subtilIs*]
53	pClosMP_53	46211	48145	1934	+	transposase [*Bacillus cereus*]
54	pClosMP_54	48169	48558	389	+	transposase [*Xanthomonas oryzae*]
55	pClosMP_55	48650	49165	515	+	transposase [*Bacillus cereus*]
56	pClosMP_56	49287	49460	173	+	hypothetical protein CdifA_16466 [*Clostridium difficile* ATCC 43255]
57	pClosMP_57	49478	49795	317	+	NONE
58	pClosMP_58	49963	50145	182	-	NONE
59	pClosMP_59	50281	50436	155	-	NONE
60	pClosMP_60	50517	50687	170	-	DNA-binding protein [*Clostridium difficile* 630]
61	pClosMP_61	51328	51642	314	+	NONE
62	pClosMP_62	51742	52104	362	+	putative membrane protein [*Stigmatella aurantiaca*]
63	pClosMP_63	52356	52481	125	+	hypothetical protein SARI_03354 [*Salmonella enterica*]
64	pClosMP_64	53233	54351	1118	+	Ubiquitin-activating enzyme E1-like [*Schizosaccharomyces sp.*]

65	pClosMP_65	54373	55785	1412	+	magnesium-protoporphyrin IX cyclase [*Syntrophobacter fumaroxidans*]
66	pClosMP_66	55998	56999	1001	+	HesA/MoeB/ThiF family protein related to EC-YgdL [*Clostridium difficile 630*]
67	pClosMP_67	57003	58688	1685	+	hypothetical protein [*Clostridium difficile 630*]
68	pClosMP_68	58708	60381	1673	+	radical SAM domaincontaining protein [*Burkholderii sp.*]
69	pClosMP_69	60414	61085	671	+	phage methyltransferase [*Staphylococcus epidermidis*]
70	pClosMP_70	61731	62162	431	+	hypothetical protein [*C. acetobutylicum*]
71	pClosMP_71	62300	62794	494	+	NONE
72	pClosMP_72	62942	63283	341	+	NONE
73	pClosMP_73	63436	63990	554	+	hypothetical protein CdifQCD-6_19248 [*Clostridium difficile*]
74	pClosMP_74	64025	64468	443	+	hypothetical protein CdifQCD-6_20423 [*Clostridium difficile*]
75	pClosMP_75	64533	65210	677	+	hypothetical DBD of Transcriptional Regulators [*Clostridium asparagiforme*]
76	pClosMP_76	65395	65787	392	+	NONE
77	pClosMP_77	65965	66204	239	+	methyltransferase domain-containing protein [*Bacteroides sp.*]
78	pClosMP_78	67222	67458	236	-	NONE
79	pClosMP_79	67501	68106	605	-	DNA-invertase BinR [*Staphylococcus aureus*]
80	pClosMP_80	68555	68770	215	+	NONE
81	pClosMP_81	69104	70534	1430	+	radical SAM domain protein [*Paenibacillus larvae subsp. larvae*]
82	pClosMP_82	70531	71766	1235	+	radical SAM domain protein [*Ruminococcus albus*]

83	pClosMP_83	71950	73587	1637	+	ABC-type transport permease component [*Thermoanaerobacter siderophilus*]
84	pClosMP_84	74070	74513	443	-	putative membrane protein [*Clostridium butyricum*]
85	pClosMP_85	74516	74716	200	-	transcriptional regulator2C Cro/CI family [*Clostridium difficile 630*]
86	pClosMP_86	75699	76316	617	+	recombinase [*Clostridium perfringens* C str. JGS1]
87	pClosMP_87	76690	77061	371	+	type IV pilus modification protein PilV partial + SH3_3 domain
88	pClosMP_88	77322	78674	1352	+	hemagglutinin-related protein, partial [*Paenibacillus dendritiformis*]
89	pClosMP_89	79178	82960	3782	-	botulinum neurotoxin [*Clostridium botulinum*]
90	pClosMP_90	83030	84505	1475	-	OrfX3 [*Clostridium botulinum*]
91	pClosMP_91	84593	86815	2222	-	OrfX2 [*Clostridium botulinum*]
92	pClosMP_92	86827	87255	428	-	OrfX1 [*Clostridium botulinum*]
93	pClosMP_93	87349	90852	3503	-	Botulinum neurotoxin type C1 nontoxic component [*Clostridium botulinum*]
94	pClosMP_94	91165	92445	1280	-	P-47 protein (Fragment) [*Paenibacillus larvae subsp. larvae* BRL-2300]
95	pClosMP_95	93211	93561	350	-	Transposase [*Bacillus sp.*]
96	pClosMP_96	94005	94385	380	-	Transposition regulatory protein TnpC [*Bacillus cereus* Rock4-2]
97	pClosMP_97	94386	96500	2114	-	Transposition regulatory protein TnpB [*Bacillus mycoides* DSM 2048]
98	pClosMP_98	96490	97623	1133	-	Transposon TnpA [*Bacillus cereus*]
99	pClosMP_99	98169	98486	317	-	NONE
100	pClosMP_100	98776	99153	377	-	Transposase [*S. aureus*]

101	pClosMP_101	99381	99815	434	-	Arginyl-tRNA synthetase [*Clostridium difficile* 630]
102	pClosMP_102	99827	100018	191	-	Electron transport complex protein rnfE [*Marinobacter aquaeolei*]
103	pClosMP_103	100022	100786	764	-	Transposase B from transposon Tn554 [*Staphylococcus aureus*]
104	pClosMP_104	100779	101879	1100	-	Transposase A from transposon Tn554 [*Staphylococcus aureus*]
105	pClosMP_105	102009	102323	314	-	DNA repair protein RadC family protein [*Clostridium difficile* 050-P50-2011]
106	pClosMP_106	102829	103029	200	-	hypothetical protein CdifQCD-6_19533 [*Clostridium difficile*]
107	pClosMP_107	104072	104197	125	-	NONE
108	pClosMP_108	104447	104833	386	-	NONE
109	pClosMP_109	104925	106238	1313	-	parB, spo0J chromosome partitioning protein [*Clostridium difficile 630*]
110	pClosMP_110	106240	106992	752	-	parA, spo0J chromosome partitioning protein [*Clostridium difficile QCD-63*]
111	pClosMP_111	107101	107298	197	-	hypothetical protein CdifQCD-6_20443 [*Clostridium difficile*]
112	pClosMP_112	107547	107798	251	-	hypothetical protein [*Acinetobacter baumannii*]
113	pClosMP_113	107995	108567	572	-	integrase family protein [*Clostridium difficile*]

+ Genes are in Forward orientation/strand on the plasmid

- Genes are in the Reverse orientation/strand on the plasmid.

The genes in the Cbp 109Kb plasmid are similar.

Table 2.2: Gene Prediction of Cbm Genome from BGI Illumina Sequencing.

Gene Number	3,835
Gene Length(bp)	3,319,296
GC Content in Gene Region(%)	28.77
Gene Length / Genome (%)	85.05
Gene Average Length(bp)	865
Intergenic Region Length(bp)	583,317
GC Content in Intergenic Region(%)	23.95
Intergenic Region Length / Genome (%)	14.94

Table 2.3. Assembly Results of Cbm from BGI Illumina Sequencing.

	Scaffolds
Total Num (#)	39
Total Length (bp)	3,902,613
Max length (bp)	1,235,731
Min length (bp)	612
Sequence GC(%)	28.05

Table 2.4: Analysis of Sequencing Data of Cb, Cbm and Cbp from BGI Data.
A. The upper panel depicts the genome sizes as ascertained by BGI Illumina sequencing data. The differences in sizes of Cb and Cbm can be explained by the presence of plasmids in Cbm.
B.The lower panel depicts the presence of eight plasmids in Cbm, two of which are present in Cb and five of which are present in Cbp.

A.

Genomic DNA	Bacterial Strains		
	CBM	CB	CBP
Size	~ 3.9 MB	~3.6 MB	~3.9 MB
Scaffolds	39	26	27
Contigs	0	13	0

B.

Plasmid Size (in kb)	Cbm Scaffold #	Bacterial Strains			
		Toxic		Non-Toxic	
		Cbm	Cbp	Cb	Cbm-mutant
1.84	30	x	x		
1.96	28	x		x	x
3.6	25	x			x
4	23	x	x		
7.2	21	x	x		
14.8	18	x	x	x	x
35.8	13	x			x
109	9	x	x		

x = presence of plasmid

Table 2.5: Mutagenesis of Cbm using gamma irradiation resulted in a loss of a number of plasmids. This loss of activity associated with plasmid loss suggests the toxin genes are present on the plasmids.

Plasmid Size (in kb)	Cbm Scaffold #	Bacterial Strains			
		Non-toxic		Toxic	
		Cbm	Cbp	Cb	Cbm-mutant
1.84	30	x	x		
1.96	28	x		x	x
3.6	25	x			x
4	23	x	x		
7.2	21	x	x		
14.8	18	x	x	x	x
35.8	13	x			x
109	9	x	x		
	x = presence of plasmid				

Authors Note:

The Cbm-mutant was generated by Dr. Nadia Qureshi using gamma irradiation.

The Cisplatin mutagenesis of Cbm experiment was designed by Dr. Nadia Qureshi.

All the genome sequencing was done by outside companies.

Chapter 3

The Cry Toxin Operon of *Clostridium bifermentans* subsp. *malaysia* Is Highly Toxic to *Aedes* Larval Mosquitoes

Abstract

The management and control of mosquito vectors of human disease currently rely primarily on chemical insecticides. However, larvicidal treatments can be effective, and if based on biological insecticides, they can also ameliorate the risk posed to human health by chemical insecticides. The aerobic bacteria *Bacillus thuringiensis* and *Lysinibacillus sphaericus* have been used for vector control for a number of decades. But a more cost-effective use would be an anaerobic bacterium because of the ease with which these can be cultured. More recently, the anaerobic bacterium *Clostridium bifermentans* subsp. *malaysia* has been re-ported to have high mosquitocidal activity, and a number of proteins were identified as potentially mosquitocidal. However, the cloned proteins showed no mosquitocidal activity. We show here that four toxins encoded by the Cry operon, Cry16A, Cry17A, Cbm17.1, and Cbm17.2, are all required for toxicity, and these toxins collectively show remarkable selectivity for *Aedes* rather than *Anopheles* mosquitoes, even though *C. bifermentans* subsp. *malaysia* is more toxic to *Anopheles*. Hence, toxins that target *Anopheles* are different from those expressed by the Cry operon.

Introduction

Vector borne diseases, specifically those transmitted by mosquitos continue to persist with constant threats of re-emergence. Aggressive insecticide usage over the years contained mosquito borne diseases, such as dengue, chikungunya and yellow fevers, filariasis and malaria, primarily to a number of tropical areas. However, these diseases began to re-surface in the 1970s and previous levels of progress in the control of these diseases could no longer be sustained (Najera, 2001). Several factors can contribute to the re-emergence of mosquito borne diseases, including vector insecticide and pathogen drug resistance, and genetic changes in both the vector and the vectored pathogens (Lederberg, 1992). Major global demographic and societal changes also contribute significantly. Moreover, with global warming and growth in the world's population, especially in the tropical regions, the incidence of mosquito borne diseases will rise. Large numbers of deaths worldwide are caused by vector borne diseases. For example, resurging epidemics of malaria are seen in areas previously thought to have been constrained (Roberts et al., 2000). Consequently mortality due to malaria is still responsible for close to a million deaths per year (Murray et al., 2012). Similarly, yellow fever causes thousands of deaths annually in the sub-Saharan Africa and tropical South America and can cause hemorrhagic fever, which is fatal 20%-50% of the time (Staples et al., 2010). In addition dengue is one of the major causes of morbidity and mortality in many Asian and South American countries (World Health Organization 2004).

The discovery of species specific biological insecticides for mosquito control produced by *Bacillus thuringiensis* serovar *israelensis* (Bti) and *Lysinibacillus sphaericus*

(Ls) has opened up new prospects for a safer and targeted mosquito control. Even though

Bti formulations have been used for more than four decades no substantial resistance has

been observed in the field, although resistance to individual Bti toxins has been observed

in the laboratory (Georghiou and Wirth, 1997). However, resistance to Ls toxins has been

observed in the field in a number of countries (Nielsen-Leroux et al., 1995; Rao et al.,

1995; Charles and Nielsen-LeRoux, 2000). This ability of organisms to adapt and

develop resistance to new insecticides indicates a need for the development or discovery

of new biological insecticides.

Clostridium bifermentans serovar *malaysia* (Cbm) is an anaerobic *Clostridium*

strain possessing larvicidal properties that was isolated in Malaysia (de Barjac et al.,

1990; Lee and Seleena, 1990a). *Clostridium* is a large and diverse genus, characterized by

Gram-positive bacteria capable of forming endospores. Most species are well known for

their metabolic properties and distinctive industrial applications. However, a small

portion is responsible for causing human and animal diseases. But only *C. bifermantans*

strains to date have shown toxicity to mosquitoes and the subsp. *malaysia* is highly toxic

to *Anopheles,* with lower toxicity to *Culex* and *Aedes aegypti* (Lee and Seleena, 1990b;

Thiery et al., 1992). Cbm cells show no toxicity to mammals and goldfish (Thiery et al.,

1992). Other studies for Cbm toxicity against mammals and various other non-target

organisms have also shown a lack of toxicity (Yiallouros et al., 1994). These

characteristics make Cbm a promising target for identification of a biological insecticide

for management of mosquito populations.

Like Bti and Ls, the toxicity Cbm is expressed during sporulation stage but decreases significantly with cell lysis (Charles et al., 1990). But, unlike Bti, Cbm produces no parasporal inclusions that could be associated with toxicity (Charles et al., 1990). Preliminary biochemical analysis suggested that three proteins of 66, 18 and 16 kDa were involved in toxicity (Nicolas et al., 1993). In addition it is possible that these proteins could aggregate into a complex, and are unstable when subjected to various methods of purification (Nicolas et al., 1993).

Subsequent efforts showed these toxic components consisted of four major proteins in crude cultures. Those putative toxins include a doublet of 66-68 kDa (*Cbm71* and *Cbm72* genes) and two other small proteins of 18 and 16 kDa (Cbm17.1 and Cbm17.2) (Nicolas et al., 1993). The *Cbm71* and *Cbm72* genes, (accession numbers X94146 and X99478, respectively) were subsequently cloned, characterized and renamed as Cry16A and Cry17A toxins, and shown to have low levels of toxicity to *Anopheles*, *Aedes* and *Culex* mosquitoes (Barloy et al., 1996; Barloy et al., 1998a). The gene encoding the Cbm17.1 and Cbm17.2 proteins was also identified (accession number Y10457) and these proteins have low amino acid similarity (44%) to a hemolysin from *Aspergillus fumigatus* (Barloy et al., 1996; Barloy et al., 1998a). Importantly all four of these proteins were immunologically unrelated toxins produced in *Bt israelensis* and *L. sphaericus,* and were determined to belong to a novel class of insecticidal toxins (Nicolas et al., 1990). Although early reports (Barloy et al., 1996; Barloy et al., 1998a) indicated the Cry16A and Cry 17A toxins had mosquitocidal activity, a subsequent report (Juarez-

Perez and Delecluse, 2001) indicated the Cry16A and Cry 17A toxins are not mosquitocidal and that the Cbm 17.1 and Cbm 17.2 proteins are not hemolytic.

In this study we show that the Cry operon encodes four genes, all of which are required as a complex for toxicity to mosquito larvae. The toxin complex, however, shows surprising species selectivity, particularly to *Aedes* mosquitoes.

Material and Methods

Bacterial strains and culture conditions

C. bifermentans (ATCC) was used as the wild type reference strain and *C. bifermentans* serovar *malaysia* (*Cb malaysia*) was from the collection of the Institute for Medical Research, Malaysia (Lee and Seleena, 1990b). These strains were used for bioassays and for isolation of gDNA for cloning. *Escherichia coli* strain DH10 beta electrocompetent cells (New England Biolabs, Ipswich, MA) were used for cloning and the NEB transformation protocol was used. *B. thuringiensis* subsp. *israelensis* 4Q7 (Bacillus Stock Center, Ohio State University, Columbus, Ohio) was used for expression of toxin genes using a published protocol (Macaluso and Mettus, 1991). Clostridial strains were streaked for isolation on TYG agar and liquid cultures were grown in TYG medium at 30°C under anaerobic conditions using BD GasPakEZ (Beckton-Dickinson Microbiology, Cockeysville, MD). The liquid cultures for *Clostridium* strains were grown anaerobically at 30°C, while the *Bacillus* cultures were grown at 30°C with shaking. Nutrient broth media was used for *B. thuringiensis* and Luria broth media was

used for *E. coli*. Ampicillin (100µg/ml) and erythromycin (25µg/ml) were added when required. To monitor mosquitocidal activities in *Cb malaysia*, the cultures after varying times of growth were used in bioassays. The cultures were also centrifuged in a high speed centrifuge to isolate pellet and supernatant fractions. *Cb malaysia* also undergo autolysis by 12 h; however, bacterial cell lysis was enhanced by sonication when needed.

General DNA Techniques

All DNA manipulations were performed according to standard protocols. Primers used in this study were created using the sequences from existing databases. Polymerase chain reactions (PCR) were performed in an automated thermacycler (C 1000 Touch ™, Bio Rad). Choice-TaqTM mastermix DNA polymerase (Denville Scientific, Metuchen, NJ) was used for all PCR reactions for products below 2kb. Advantage 2x polymerase (Clontech, Mountain View, CA) or Phusion polymerase (NEB) was used for all PCR reactions for products above 2kb. PCR products were separated in 1% agarose gels and subsequently cut and purified using Wizard SV Gel and PCR purification kits (Promega, Madison, WI). Sequencing of purified DNA products was performed by the Genomics Core facility at the University of California, Riverside.

Construction of plasmids expressing the full and partial Cry operon and individual toxins

The vector pHT315 (Arantes and Lereclus, 1991) was used for construction of expression plasmids (Table 3.1). Sub-cloning was accomplished using a Gibson assembly

mastermix (NEB). The Cyt1A promoter a constitutive promoter from *Bt israelensis* was used for expression of all gene constructs. Hence 508bp upstream of the first ATG of *Cyt1A*, was amplified from pWF45 (Wu and Federici, 1993) using primers 1 and 2 (Table3.2). The cassette containing the genes of interest were amplified using primers 3 and 4 for pCryO (Cry operon), primers 3 and 5 for pCry16 (*Cry16A*) and primers 4 and 6 for pCbm17.2 (*Cbm17.2*) (Table 3.2). pHT315 was linearized with SmaI in the multiple cloning site. All DNA products were gel purified and purified using Wizard SV Gel and PCR (Promega) purification kits to provide the desired DNA concentrations. A Gibson assembly protocol (Gibson et al., 2009) was used to assemble the fragments, followed by transformation in DH10 beta electro-competent cells (NEB). Plasmids, pCryO (Cry operon), pCry16 (*Cry16A*) and pCbm17.2 (*Cbm17.2*), were made using this approach. However, for plasmids, pCry17 (*Cry17A*) and pCbm17.1 (*Cbm17.1*), the genes were synthesized (Genscript, Piscataway, NJ) and cloned in pHT315. The constructs pCry16/17 and pCryO/Δ17.1 were obtained using *Nco*I and *Pac*I restriction, respectively, of pCryO followed by religation of the resulting restricted plasmid. Following confirmation of correct ligation, the plasmid constructs (Table 3.1) were extracted from DH10 beta cells in ample quantities. All plasmids were then used to independently transform *Bt israelensis* 4Q7 cells by electroporation and colonies isolated for erythromycin resistance as previously described (Chang et al., 1993).

Bioassays

Total cultures of *C. bifermantans*, *Cb malaysia* and recombinant *B. thuringiensis* (rBT) strains were assayed against *Ae. aegypti, An. gambiae An. stephensi*, and *C. quinquefasciatus* larvae. Bioassays were performed at 24° C using 25 third instar larvae in 200 ml water. Live bacterial cultures were used in bioassays. Larval mortality was determined by counting the number of larvae still alive at 24, 48 or 72 h. Bioassays were repeated 2-3 times and LC50 concentrations were determined by probit analysis (USDA) and plotted using the Origin program (Origin Lab, Northampton, MA).

Hemolytic Assays

One ml of trypsinized sheep red blood cells (RBC) (Colorado Serum Company) were centrifuged at 2000xg for 2 min. The pellet was washed three times with ice cold 1X PBS (pH7.4). The pellet was then resuspended in 30 ml 1X PBS (pH 7.4). To 900 μl of the RBC suspension, 100 μl of the bacterial cell culture was added and the mixture incubated at 37°C for 30 min. The O.D 600 was measured using a spectrophotomer (SmartSpec 3000, BioRad, Hercules, CA). RBC suspension and 1X PBS(pH 7.4) were used as negative controls and 0.1% Triton and Cyt1A (15 ng/ml) were used as positive controls. Whole culture samples and supernatant from B-PER® Bacterial Protein Extraction Reagent (Pierce, Rockford, IL) of rBT strains containing pCryO, pCbm17.1, pCbm17.2, and CryO/Δ17.1 were used as samples.

Antibodies

Hydrophilic peptides from the Cry16A and Cbm17.1 were designed based on TopPred topology and antigenicity prediction index (http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html). The Cry16A peptide (CSYTDGNFEDFPKLS) comprising residues 591–604 and Cbm17.1 peptide (CNKLTIDKYNTKFAI) comprising residues 115-128, were commercially synthesized, confirmed by mass spectral analyses (GenScript USA, Piscataway, NJ), and conjugated via the introduced cysteine to a maleimide-activated KLH carrier protein (Pierce, Rockford, IL) according to the manufacturer's protocol. This conjugate was used to immunize rabbits. Cbm17.1 is comprised of 153 amino acids, and shows 78 percent identity to Cbm17.2, which has amino acids. The Cbm17.1 peptide used for antibody development is in a divergent region between the two proteins, and hence this antibody does not detect the Cbm17.2 protein.

Polyacylamide gel electrophoresis (PAGE) and immunoblot

C. bifermantans, *Cb malaysia*, rBT and 4Q7 strains were grown in parallel for 4-72 h. Aliquots, 100ul, of the cultures were mixed with of 2X sample buffer (Bio-Rad, Hercules, CA), and the samples separated by SDS-polyacrylamide (10-14%) gel electrophoresis. Native polyacrylamide (10%) gel electrophoresis was performed using 1X native PAGE buffer (3g Tris, 14.5g glycine). In parallel the separated proteins were transferred onto PVDF membranes (Millipore, Billerica, MA). The membranes were treated with blocking solution (1X PBS, 5% skim milk and 0.1% Tween-20) for 1 h at

room temperature and then washed with PBST (1XPBS and 0.1% Tween-20). The

blocked membrane was incubated overnight (at 4°C) with primary anti-Cry16A or

Cbm17.1 antibody at 1:1000 dilution. A secondary antibody, ECL™ anti-rabbit IgG-

horseradish peroxidase-linked whole antibody (GE Healthcare, Anaheim, CA) was used

at a dilution of 1:5000. Immunoreactive bands were visualized using the ECL western

blotting kit (GE Healthcare) and exposed to an X-ray film.

Mass Spectrometry.

Since there were substantial differences in mosquitocidal activity between 1- and

5-day cultures, these bacterial cultures were separated by SDS-PAGE. The protein bands

that differed between the two cultures were excised from the gel, digested by trypsin, and

analyzed by nano-ultra performance liquid chromatography/tandem mass spectrometry

(nano-UPLC/MS/MS) at the IIGB proteomics facility, UC Riverside.

Results

Toxicity of Clostridium bifermantans malaysia to mosquito larvae.

In this study we used a proteomics approach to identify the *Cb malaysia* toxins

involved in conferring toxicity to towards mosquito larvae. To facilitate such analysis we

analyzed proteins that were active in the pellet and supernatant fractions obtained after

centrifugation of the bacterial cultures. Thus we monitored the toxicity of *Cb malaysia*

live cultures to 3rd or 4th instar *An. stephensi* mosquito larvae. At 6-10 h, a high level of

118

toxicity was observed in the bacterial pellet while no toxicity was observed in the supernatant. However, at 12-24 h high levels of toxicity were also observed in supernatant fractions, as well as the pellet fraction. The biological activity remains high till 72 h (Figure 3.1).

To determine the stability of the mosquitocidal activity, cells were lysed and tested for toxin stability after incubation at different temperatures. Under the tested conditions, it appeared that following 6 h incubation at -80°C, 0°C and 8°C toxicity remained high, but decreased at higher temperatures. However, the mosquitocidal activity was lower at all temperatures with the complete loss of toxicity at 55°C after 24 h incubation. The data shows that the overall half-life of the toxicity was about 1 day and the toxin was unstable at high temperatures. The loss of activity is likely a combination of proteolysis and low stability of *Cb malaysia* toxins.

Mosquitocidal activity of supernatants obtained after cell lysis was also determined at 1- and 5-day. At 1-day culture a high level of toxicity was observed in the supernatant fraction while the 5-day culture showed no toxicity, suggesting that the toxins are degraded with time after cell lysis (Figure 3.2). Hence proteins from 1- and 5-day cultures were separated immediately by SDS PAGE and proteins that differed were subjected to analysis by UPLC/MS/MS. Among the 21 proteins identified in day 1 band A, 11 of these were not present on day 5. Of these four were DNA gyrase, a peptidase, an acetate kinase, a conserved protein, two hypothetical proteins and the Cry16A toxin.

Expression of the Cry toxin operon

The Cry16A toxin as reported previously is encoded by an operon (Barloy et al., 1996). Hence the full operon (Figure 3.3A, Table 3.1) consisting of the *Cry16A, Cry17A, Cbm17.1* and *Cbm17.2* genes was expressed using the Cyt1A promoter from *Bt israelensis* (Chang et al., 1993; Wu and Federici, 1993).

Protein expression from this transformation, rBT pCryO, was analyzed by Coomassie Blue staining and western blotting (Fig. 3B). Cultures were grown for 4, 12, 18, 24 and 48 h to determine the optimal time for toxin production. A strong band was observed around 70kDa for the Cry16A and Cry17A, which was visible in 48 h cultures, but the approximately 20 kDa band for Cbm17.1 and Cbm17.2, which was visible at 4 h, was not visible by around 48 h of culture. These bands were not present in the control comprising of 4Q7 strain containing the pHT315 vector alone. Thus we concluded the entire operon is expressed under these conditions, with an optimal expression of the operon at around 24 h.

Western blotting against whole cultures with antibodies raised against the Cry16A and Cbm17.1 peptides also revealed the presence of a protein at about 70 kDa corresponding to the predicted size of the Cry16A toxin, and one at about 20kDa, which corresponds to the predicted size of the Cbm17.1 and Cbm17.2 proteins. Both of these bands could be detected in bacterial strains containing pCryO at 24 h. Moreover, the proteins reacting with Cry16A antibodies were observed in whole cultures of *Cb malaysia* (Figure 3.3C), but not in *C. bifermantans* cultures. No protein reacting with

Cry16A or Cbm17.1 antibodies was observed in whole cultures of *C. bifermantans* or the

4Q7 containing the pHT315 vector alone (Figure 3.3C).

Larvicidal activity of bacterial clone expressing the Cry operon.

Whole cell preparations from *C. bifermantans*, *Cb malaysia*, and rBT

transformants containing pHT315 (control) and pCryO that expressed the entire Cry

operon were assayed for larvicidal activity using third instar *Ae. aegypti*, *An. gambiae*

(Table 3) and *C. quinquefasciatus*. The rBT pHT315 control samples showed no larval

toxicity against all three species. But the rBT pCryO showed no activity against *An.*

gambiae even though *Cb malaysia* is highly toxic to this species with an LC50 value of

6.4 x 105 cfu/ml (Figure 3.4A, Table 3.3). In contrast the rBT pCryO showed high

larvicidal activity against *Ae. aegypti* with a pattern comparable to that of *Cb malaysia*

but showing higher mortality (Figure 3.4B). *Cb malaysia* had an LC50 value of 3.3 x 108

cfu/ml, whereas rBT pCryO had an LC50 value of 1.4 x 108 cfu/ml (Figure 3.4B, Table

3.3). rBT pCryO also showed no toxicity to *C. quinquefasciatus* larvae.

Larvicidal activity of individual Cry operon toxins.

Since the entire Cry operon that expressed all four of the toxins showed high

toxicity to *Aedes* larvae, we then analyzed the role of the individual toxins. These toxins

were cloned under control of the Cyt1A promoter and the Cry operon terminator, giving

pCry16, pCry17, pCbm17.1 and pCbm17.2 (Figure 3.5A). Of these only the pCry16A

construct was observed to form a stable protein (Figure 3.5B). The Cry17A protein was

expressed but readily degrades in *Bacillus* cells even after 24 h, while the Cbm17.1 and Cbm17.2 proteins were poorly expressed on their own. Importantly none of the bacterial strains containing these constructs showed any larval toxicity (Figure 3.5C).

Based on these data we concluded that some of these toxins could be acting together. We then made deletion constructs using *Pac*I and *Nco*I restriction of the Cry operon. Restriction of pCryO and followed by ligation of the restricted fragment led to the construction of the plasmids pCryO/Δ17.1 (deletion of *Pac*I fragment) and pCry16/17 (deletion of *Nco*I fragments) (Figure 3.6A). The pCryO/Δ17.1 and pCty16/17 constructs expressed protein(s) of 70kDa (Figure 3.6B), and the Cry16A protein was detected using an anti-Cry16A antibody. However, both of these constructs were not toxic to *Aedes* larvae.

The Cry operon toxins form a complex

Since none of the constructs made showed any toxicity, other than pCryO that includes all proteins in the operon, we hypothesized that a complex may be needed for toxicity. Indeed analysis by native PAGE showed a complex of about 170kDa (Figure 3.7A). In contrast, rBT containing pCry16, pCry17 or pCryO/Δ17.1 did not show the formation of a significant amount of the complex. The complex formed cross-reacted with the anti- Cbm17.1 antibody (Figure 3.7B). Less distinct and larger complexes are also visible in both native PAGE and immunoblots of bacterial cultures expressing the entire Cry operon (Figure 3.7A, B).

The Cry operon lacks hemolytic activity

Hemolytic activity rBT strains containing pCryO, pCbm17.1, pCbm17.2, and CryO/Δ17.1 was analyzed using sheep RBC. None of these strains showed any hemolytic activity in either the pellet or B-Per bacterial extracts. In contrast, both 0.1% Triton and Cyt1A (15 ng/ml) effectively lysed the RBC.

Discussion

The novel bacterial strain *C. bifermantans malaysia* is highly active against a number of mosquito genera but particularly against anopheline species (de Barjac et al., 1990; Lee and Seleena, 1990b). However, the genes responsible for its toxicity have remained elusive. Prior efforts characterized the bacteria and identified proteins that had mosquitocidal activity (Charles et al., 1990; Nicolas et al., 1990; Nicolas et al., 1993), and the genes encoding these proteins were subsequently identified (Barloy et al., 1996; Barloy et al., 1998b). All of the identified proteins (Nicolas et al., 1993) are encoded in a single operon, the Cry operon (Barloy et al., 1996).

This operon consists of four genes, *Cry16A, Cry17A, Cbm17.1* and *Cbm17.2*. The derived sequence of Cry16A and 17A suggests that these proteins are part of Cry-like toxin family and are similar in size at about 70kDa (Barloy et al., 1996). The Cry17A has about 30% amino acid identity to the Cry27A and Cry29A proteins, while the Cry16A has about 30% identity to the Cry19A protein (http://www.btnomenclature.info). In contrast the Cbm17.1 and 17.2 proteins belong to the aegerolysin family of proteins

123

(Berne et. al., 2009). These two proteins are 78% similar in amino acid sequence and are 17kDa in size (Barloy et al., 1996; Barloy et al., 1998a, b) but run as 20 kDa proteins in SDS-PAGE.

Expression of the Cry operon showed low levels of activity to *Anopheles, Aedes* and *Culex* larvae (Barloy et al., 1996). Importantly, in this study expression of only the Cry16A protein and part of the Cry17A protein also showed toxicity to all three mosquito species. However, subsequent efforts from this same group suggested that all these proteins are neither mosquitocidal or hemolytic (Barloy et al., 1998b; Juarez-Perez and Delecluse, 2001).

In our efforts to identify novel mosquitocidal genes, we reinvestigated toxins in this bacterial strain. Proteomics of SDS-PAGE bands of this strain again identified the Cry16A and Cbm17 proteins as potential toxins, since these were observed in 1 day cultures that were toxic, but not in 5 day cultures that were in active (Figure 3.2B). The Cry operon encodes four proteins but the individual proteins did not appear to have toxicity (Juarez-Perez and Delecluse, 2001). The presence of a single promoter upstream of four genes, each with intact translation sites suggested (Barloy et al., 1996) that these genes were expressed simultaneously. Based on this premise we created and expressed a clone of the entire operon under the control of a strong *Bacillus* promoter, Cyt1A (Chang et al., 1993; Wu and Federici, 1993).

Expression of the 70 kDa Cry proteins is clearly visible at 48 and 72 h. However, western blot shows the protein can be observed at least by 24 h, and likely earlier. This may be because as previously observed, very small amounts of Cry16A is secreted before

sporulation (Barloy et al., 1998b). The Cbm17.1 and Cbm17.2 proteins can be observed in as early as 12 h cultures. Since both the Cry16A and Cbm17 proteins are expressed at 24 h, we routinely used such cultures for bioassays.

We show here that when the Cbm Cry operon was expressed in its entirety, high levels of larvicidal activity towards *Aedes aegypti* was observed, and the toxicity of the 4Q7 *Bacillus* strain expressing the Cry operon, rBT_CryO, is higher than that observed with wild type Cbm (Figure 3.4B). However, to our surprise this 4Q7 *Bacillus* strain expressing the Cry operon showed no toxicity to *An. gambiae* and *C. quinquefasciatus* larvae. Our observations differ from those of Barloy et al., (1996) since we observed much higher levels of toxicity of the Cry operon to *Aedes* mosquitoes. Potentially the use of a stronger *Bacillus thuringiensis* promoter, like Cyt1A, results in the production of high toxin levels observed here. Further we did not observe any toxicity to *Anopheles* or *Culex* mosquitoes, nor was any of the truncated Cry operon constructs active as previously observed (Barloy et al., 1996).

Since the Cry operon was larvicidal the four genes were individually expressed under control of the Cyt1A promoter. None of these proteins were toxic, although we could confirm only the expression of the full-length protein for only one toxin, Cry16A. Consequently constructs that had deletion of the *Cbm17.1* gene or both the *Cbm17* genes were made, but both were also inactive. Thus our systematic deletion of genes from the original active clone, pHT315_CryO, supports the original hypothesis that all genes within the Cry operon are required for larvicidal activity. Our results are therefore

consistent with previous observations (Nicolas et al., 1993) that a toxin complex is required for toxicity.

Cbm is the first known anaerobic bacterium possessing larvicidal activity. The Cry larvicidal genes are similar to those found in *B. thuringiensis*. Despite the similarities Cry16A did not cross-react with antibodies raised against any of the *B. thuringiensis* subsp. *israelensis* or other *B. thuringiensis* toxins (Nicolas et al., 1993). Further the hemolysin like proteins, Cbm17.1 and Cbm17.2, are similar to proteins found in various fungi including *Aspergillus*, but are not identical. However, the Cry operon does not show any hemolytic activity. Potentially this is because the toxins are indeed non-hemolytic or alternatively, the Cbm17 toxins were not obtained in an active conformation when isolated. Nevertheless it is clear the Cry operon in *Cb malaysia* contains a very unique combination of toxins that act cooperatively. It is a likely possibility that Cbm acquired this operon with horizontal gene transfer.

The Cry operon shows high selectivity towards *Aedes* mosquitoes. In contrast the wild type Cbm shows much higher toxicity to *An. gambiae* than to *Ae. aegypti*. Consequently, the Cbm proteins that are responsible for *Anopheles* toxicity are likely to be different and are unknown, and investigations are underway to identify these. In addition, in future studies, we plan to investigate the mechanisms of larvicidal activity of the Cry toxins and the role these proteins play in the toxicity towards *Aedes* larvae.

References

Arantes, O. and Lereclus, D. (1991). Construction of cloning vectors for *Bacillus thuringiensis*. *Gene*. **108**: 115-119.

Barloy, F., Delecluse, A., Nicolas, L., and Lecadet, M.M. (1996). Cloning and expression of the first anaerobic toxin gene from *Clostridium bifermentans* subsp. *malaysia*, encoding a new mosquitocidal protein with homologies to *Bacillus thuringiensis* delta-endotoxins. *J Bacteriol*. **78**: 3099-3105.

Barloy, F., Lecadet, M.M., and Delecluse, A. (1998a). Distribution of clostridial cry-like genes among *Bacillus thuringiensis* and *Clostridium* strains. *Curr Microbiol*. **36**: 232-237.

Barloy, F., Lecadet, M.M., and Delecluse, A. (1998b) Cloning and sequencing of three new putative toxin genes from *Clostridium bifermentans* CH18. *Gene*. **211**: 293-299.

Berne S, Lah L, Sepcic K. (2009). Aegerolysins: structure, function, and putative biological role. *Protein Sci*. **18**:694-706.

Chang, C., Yu, Y.M., Dai, S.M., Law, S.K., and Gill, S.S. (1993). High-level cryIVD and cytA gene expression in *Bacillus thuringiensis* does not require the 20-kilodalton protein, and the coexpressed gene products are synergistic in their toxicity to mosquitoes. *Appl Environ Microbiol*. **59**: 815-821.

Charles, J.-F., and Nielsen-LeRoux, C. (2000). Mosquitocidal bacterial toxins: Diversity, mode of action and resistance phenomena. *Mem Institut Oswaldo Cruz*. **95**: 201-206.

Charles, J.F., Nicolas, L., Sebald, M., and de Barjac, H. (1990). *Clostridium bifermentans* serovar *malaysia*: sporulation, biogenesis of inclusion bodies and larvicidal effect on mosquito. *Research Microbiol*. **141**: 721-733.

de Barjac, H., Sebald, M., Charles, J.F., Cheong, W.H., and Lee, H.L. (1990). *Clostridium bifermentans* serovar *malaysia*, a new anaerobic bacterium pathogen to mosquito and blackfly larvae. *C R Acad Sci III*. **310**: 383-387.

Georghiou, G.P., and Wirth, M.C. (1997). Influence of exposure to single versus multiple toxins of *Bacillus thuringiensis* subsp. *israelensis* on development of resistance in the mosquito *Culex quinquefasciatus* (Diptera: Culicidae). *Appl Environ Microbiol*. **63**: 1095-1101.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*. **6**: 343-345.

Juarez-Perez, V., and Delecluse, A. (2001). The Cry toxins and the putative hemolysins of *Clostridium bifermentans* serovar *malaysia* are not involved in mosquitocidal activity. *J Invertebr Pathol*. **78**: 57-58.

Lederberg, J., Shope R.E., Oakes S.C. (1992). Emerging infections: microbial threats to health in the United States. *Washington: National Academy Press*

Lee, H.L., and Seleena, P. (1990a). Isolation of indigenous larvicidal microbial control agents of mosquitos: the Malaysian experience. *Southeast Asian J Trop Med Public Health*. **21**: 281- 287.

Lee, H.L., and Seleena, P. (1990b). Isolation and evaluation of larvicidal *Clostridium bifermantans* against mosquitoes of public health importance. *Trop Biomed*. **7**: 103-106.

Macaluso, A., and Mettus, A.M. (1991). Efficient transformation of Bacillus thuringiensis requires nonmethylated plasmid DNA. *J Bacteriol.* **173**: 1353-1356.

Murray, C.J., Rosenfeld, L.C., Lim, S.S., Andrews, K.G., Foreman, K.J., Haring, D. (2012). Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet.* **379**: 413- 431.

Murray NE, Quam MB, Wilder-Smith A. (2013). Epidemiology of dengue: past, present and future prospects. Clin Epidemiol **5:**299-309.

Najera, J.A. (2001). Malaria control: achievements, problems and strategies. *Parassitologia*. **43**: 346 1-89.

Nicolas, L., Hamon, S., Frachon, E., Sebald, M., and Barjac, H. (1990). Partial inactivation of the mosquitocidal activity of *Clostridium bifermentans* serovar *malaysia* by extracellular proteinases. *Applied Microbiol. Biotechnol.* **34**: 36-41.

Nicolas, L., Charles, J.F., and de Barjac, H. (1993). *Clostridium bifermentans* serovar *malaysia*: characterization of putative mosquito larvicidal proteins. *FEMS Microbiol Lett.* **113**: 23-28.

Nielsen-Leroux, C., Charles, J.F., Thiery, I., and Georghiou, G.P. (1995). Resistance in a laboratory population of *Culex quinquefasciatus* (Diptera: Culicidae) to *Bacillus sphaericus* binary toxin is due to a change in the receptor on midgut brush-border membranes. *Eur J Biochem*. **228**: 206-210.

Rao, D.R., Mani, T.R., Rajendran, R., Joseph, A.S., Gajanana, A., and Reuben, R. (1995). Development of a high level of resistance to *Bacillus sphaericus* in a field population of *Culex quinquefasciatus* from Kochi, India. *J Am Mosq Control Assoc*. **11**: 1-5.

Roberts, D.R., Manguin, S., and Mouchet, J. (2000). DDT house spraying and re-emerging malaria. *Lancet.* **356**: 330-332.

Staples, J.E., Gershman, M., and Fischer, M. (2010). Yellow fever vaccine: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recommendations and reports : Morbidity and mortality weekly report Recommendations and reports. *Centers for Disease Contro.* **59**: 1-27.

Thiery, I., Hamon, S., Gaven, B., and De Barjac, H. (1992) Host range of *Clostridium bifermentans* serovar *malaysia*, a mosquitocidal anaerobic bacterium. *J Am Mosq Control Assoc*. **8**: 272-277.

World Health Organization. (2004). World Health Report. WHO, Geneva.

Wu, D., and Federici, B.A. (1993) A 20-kilodalton protein preserves cell viability and promote CytA crystal formation during sporulation in *Bacillus thuringiensis*. *J Bacteriol.* **175**: 5276- 5280.

Yiallouros, M., Storch, V., Thiery, I., and Becker, N. (1994). Efficacy of *Clostridium bifermentans* serovar *malaysia* on target and nontarget organisms. *J Am Mosq Control Assoc*. **10**: 51-55.

Figure 3.1. Toxicity appears in supernatants within a day of *C. bifermantans malaysia* cultures. To analyze the appearance of toxicity in the supernatants, both the pellet and supernatant fractions were monitored for toxicity to fourth instar *An. stephensi* larvae. A constant amount of the bacterial culture was assayed. High level of toxicity in the pellet of the bacterial culture was observed within a short time of culture, while no toxicity was detected in the supernatant until 12 h when lysis of sporangia occurs(16). Toxicity was monitored at 24, 48 and 72 h.
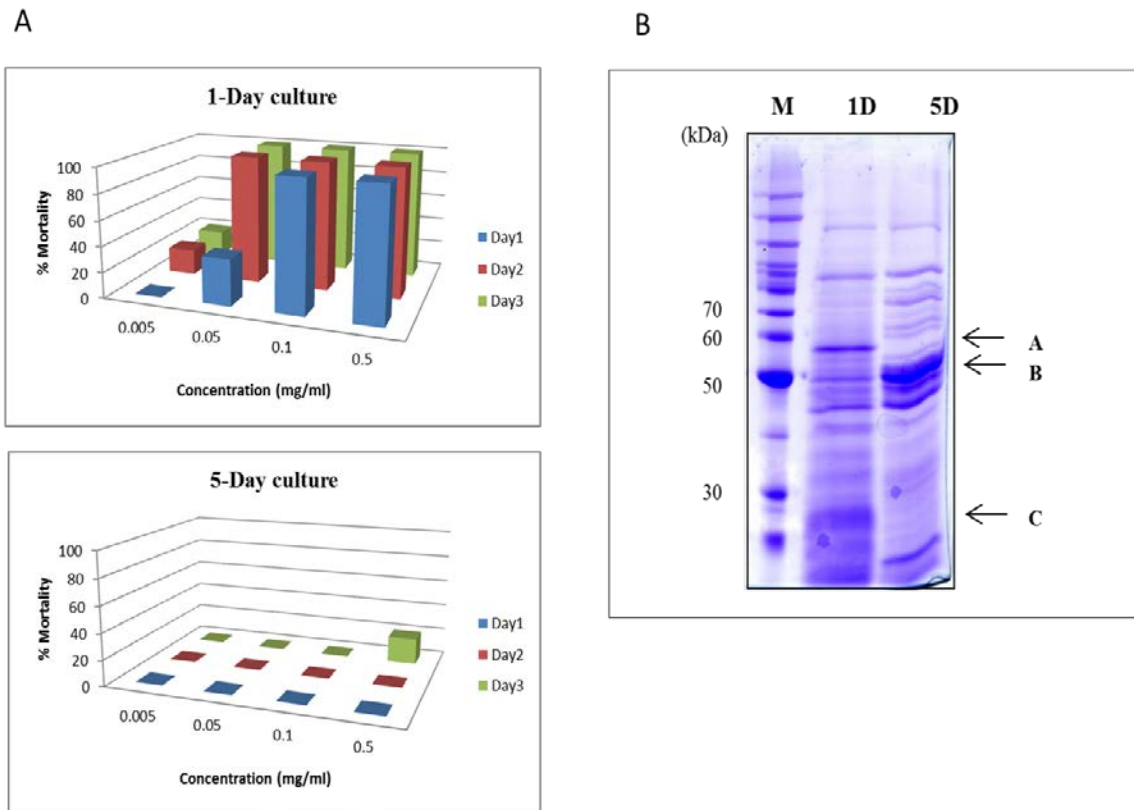
Figure 3.2. Toxicity is lost in *C. bifermantans malaysia* cultures after five days.

A. **Toxicity in one and five day supernatants.** Whole cultures were centrifuged to collect supernatants. At one day culture a high level of toxicity was observed in the supernatant. However, at five days no toxicity was observed in in 24 and 48 h bioassays.

B. **Differential expressions of the protein profiles for one and 5 day cultures were observed in the supernatants (band A and B).** Following tryptic digestion the peptides were sequenced using tandem mass spectrophotometry. Band B of five day supernatant was identified as heat shock proteins. In contrast, band A was found in 1 day but not in 5 day supernatants. One protein identified by proteomics was the Cry16A protein in the 1 day supernatant.

131
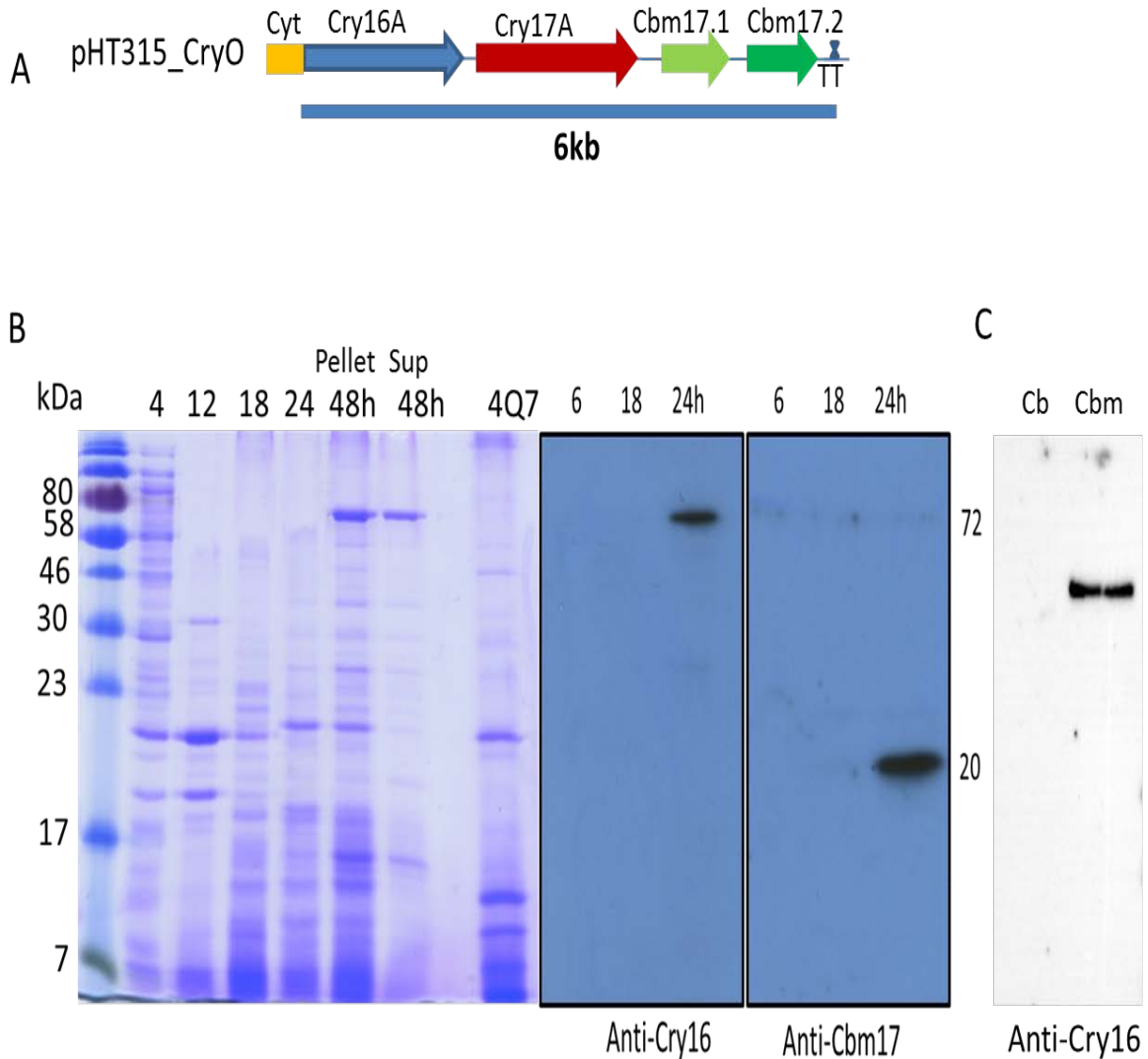
Figure 3.3. The full-length operon expresses both the Cry and hemolysin proteins.

A. **Schematic of the Cry operon under control of the Cyt1A promoter from Bt israelensis.** The operon consists of the *Cry16A, Cry17A, Cbm17.1* and *Cbm17.2* genes, and a terminator (TT). Except for *Cry16A* the native Shine-Dalgarno (SD) sequences were used. For *Cry16A* the Cyt1A SD was used.

B. **Both the Cry16A and Cbm17.1 proteins were expressed when the entire operon was used for expression.** Immunoblots using anti-Cry16A and Cbm17.1 antibodies show these were expressed at 24 h.

C. The Cry16A protein was also observed in Cbm cultures at 18 h but not in Cb cultures.

132
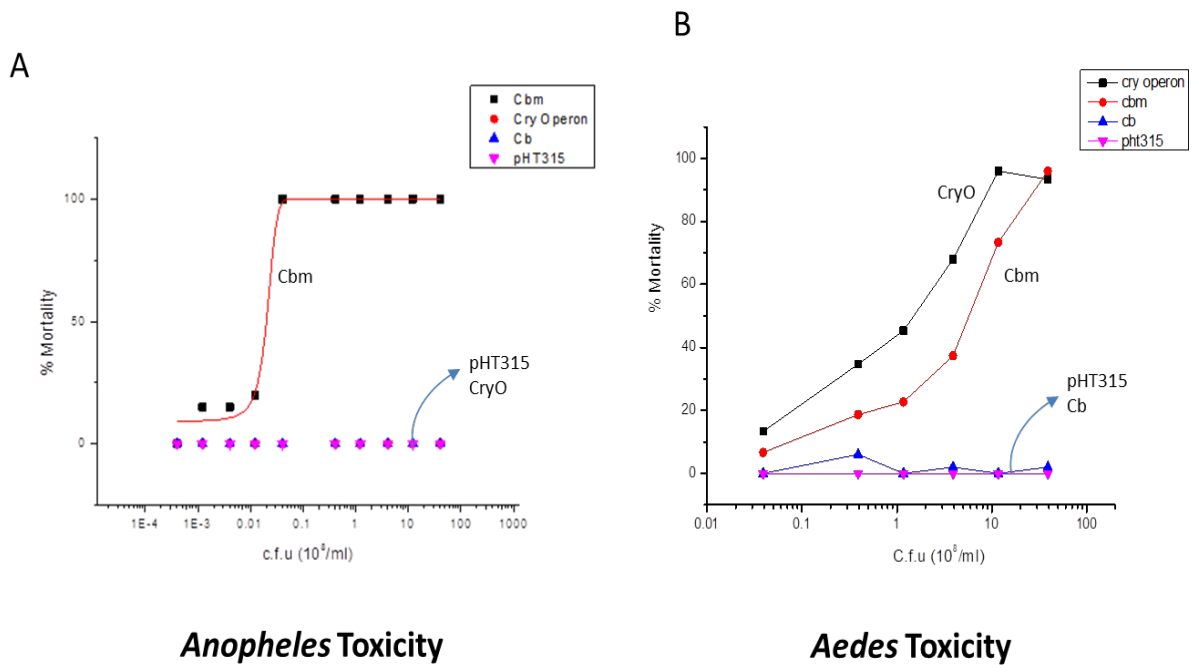
Figure 3.4. The full-length Cry operon shows high toxicity to *Aedes* mosquitoes but not to *Anopheles* mosquitoes.

- *A.* **Toxicity of Cbm and the Cry operon to Anopheles mosquitoes.** While Cbm cultures show high toxicity to *An. gambiae* mosquitoes, the Cry operon is non-toxic to both *An. gambiae* mosquitoes. The toxicity curves for the Cry operon with *An. stephensi* and *C. quinquefasciatus* were similar to those obtained with *An. gambiae*.
- *B.* **Toxicity of Cbm and the Cry operon to Aedes mosquitoes.** Both Cbm and the rBt_Cry operon were toxic to *Aedes* mosquitoes. However, Cb and the rBT_pHT315 were nontoxic.

133
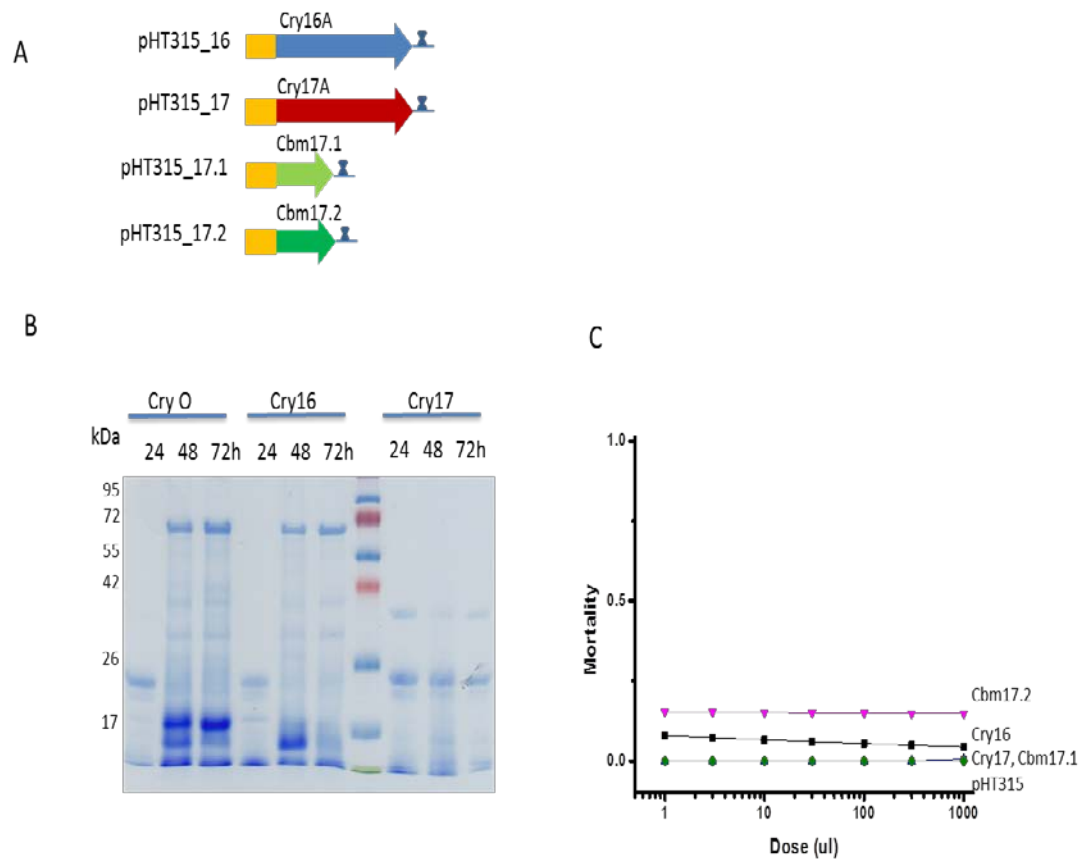
Figure 3.5. All single genes show no toxicity to *Aedes* mosquitoes

- ***A.*** Each of the toxin genes in the Cry operon was expressed under the control of the Cyt1A promoter, the Cyt1A Shine Dalgarno sequence and includes a 3 prime terminator identical to the Cry operon construct in Fig. 3.
- ***B.*** **The Cry16A toxin was stably expressed but the Cry17A is expressed but readily degrades in culture even at 24h**. The Cry16A protein is not observed in SDS-PAGE at 24 h (see also Fig. 3) but is detected by immunoblot at this time (see Fig. 3).
- ***C.*** None of the single gene constructs expressed in Bt showed any toxicity.

134
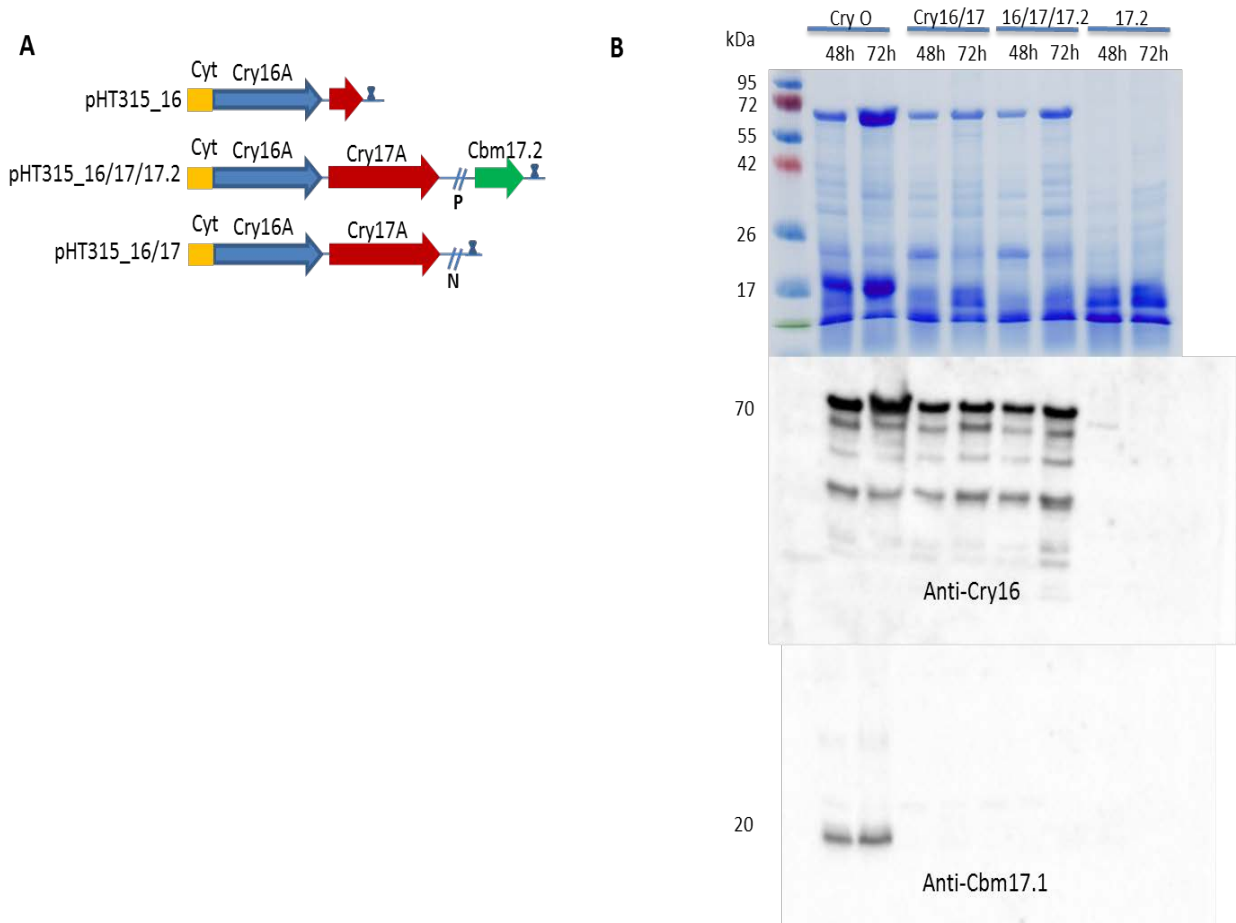
Figure 3.6. Any gene deletion in the Cry operon results in loss of toxicity.

A. **Three gene deletions were made**. In the first, pCry16, only Cry16A was expressed (see Fig. 5), while in the second, CryO/Δ17.1, only *Cbm17.1* was deleted and in the third, pCry16/17, both *Cbm17* genes were deleted. The latter two were derived from the pCryO by restriction with *Pac*I (P) and *Nco*I (N), respectively. Each of the constructs is under the control of the Cyt1A promoter and includes a 3 prime terminator identical to the Cry operon construct in Fig. 3.

B. The Cry16A toxin was observed in the rBT-CryO, rBT-p CryO/Δ17.1 and rBT-pCry16/17 cultures. The rBT-CryO construct also shows expression of the Cbm17.1 protein at 48-72 h, as observed at 24h (Fig. 3)

135
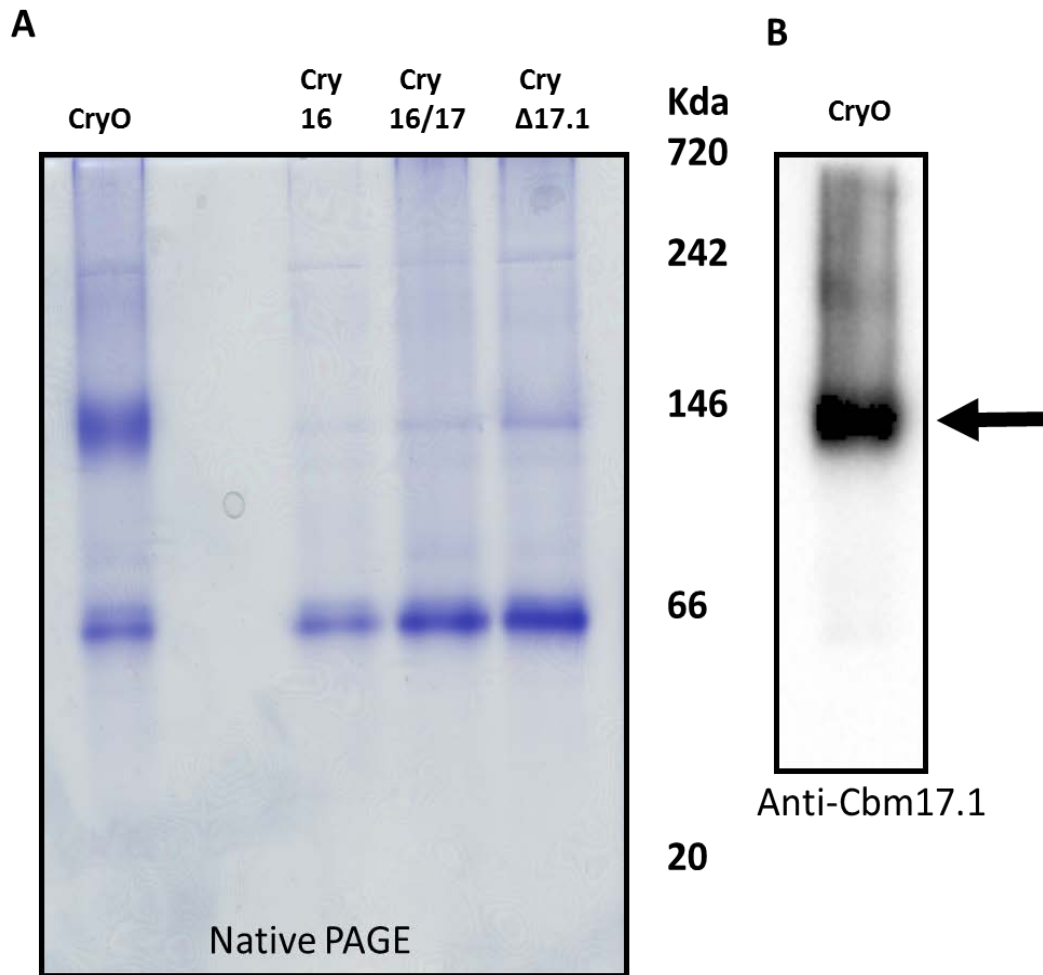
Figure 3.7. Native PAGE shows the Cry operon forms a complex.

A. The entire Cry operon (CryO) forms a complex of about 170kDa in 24 h rBTcultures. This complex included the Cbm17.1 protein as shown in the immunoblot in B.

B. None of the other three constructs, pCry16, pCry16/17 and p CryO/Δ17.1 showed the presence of the complex in rBT cultures.

Table 3.1. List of constructs used to assess toxicity of genes in the Cry operon

Construct Name[a]	*Cb malaysia* Genes Used
pCryO	*Cry16A, Cry17A, Cbm17.1, Cbm17.2*
pCry16	*Cry16A*
pCry17	*Cry17A*
pCbm17.1	*Cbm17.1*
pCbm17.2	*Cbm17.2*
pCry16/17	*Cry16A, Cry17A*
pCryO/Δ17.1	*Cry16A, Cry17A, Cbm17.2*

[a] All constructs were in pHT315 using the Cyt1A promoter from *B. thuringiensis israelensis*

Table 3.2. Primers used to create pCryO, pCry16 and pCbm17.2.

#	Primer description	Sequence
1	Promoter_f_overlap_pHT315	AGTGAATTCGAGCTCGGTACCCTTTTCGATTTCAAATTTTCCAAAC
2	Promoter_r_no_overlap	AAATAAACAACTCCTTAAAGTTAATTAGAATAGTGG
3	CryO_f_overlap_promoter	CTAATTAACTTAAGGAGTTGTTTATTTATGAATACTAATATTTTTCTACACATG
4	CryO_r_overlap_pHT315	GGTCGACTCTAGAGGATCCCCCATCTGTATCATCTCATAAAAGATTAGGG
5	Cry16A_r_overlap_pHT315	GGTCGACTCTAGAGGATCCCCAAGATTATCATTTGAGCTTGTGCCAG
6	17.2_f_overlap_promoter	CTAATTAACTTAAGGAGTTGTTTATTTATGAACAATAATTGTGAAGTTAATTGTGAG

Table 3.3. Bioassay data against *Aedes aegypti* used in Figure 3.4.B. (a)Cry operon and (d)pHT315 cultures were grown for 24h at 30°C with shaking. (b)Cbm and (c)Cb cultures were grown in anaerobic jar for 18h at 30°C with no shaking. 25 third-instar larvae put in 200ml water in plastic cups were used for bioassay. %mortality was recorded after 24 hr.

(a)

Cry operon								
Dose (in µl)	Replicate 1		Replicate 2		Replicate 3		Plotted Data	
	cfu/ml	%mortality	cfu	%mortality	cfu	%mortality	cfu/ml-ave	%mortality-ave
1	0.05664	40	0.04553	0	0.05664	0	0.0529367	13.33333333
10	0.5664	84	0.4553	20	0.5664	0	0.5293667	34.66666667
30	1.6992	100	1.3659	28	1.6992	8	1.5881	45.33333333
100	5.664	100	4.553	48	5.664	56	5.2936667	68
300	16.992	100	13.659	100	16.992	88	15.881	96
1000	56.64	84	45.53	96	56.64	100	52.936667	93.33333333

(b)

Cbm								
Dose (in µl)	Replicate 1		Replicate 2		Replicate 3		Plotted Data	
	cfu/ml	%mortality	cfu	%mortality	cfu	%mortality	cfu/ml-ave	%mortality-ave
1	0.03594	0	0.0407	8	0.0407	12	0.0391133	6.666666667
10	0.3594	8	0.407	8	0.407	40	0.3911333	18.66666667
30	1.0782	8	1.221	28	1.221	32	1.1734	22.66666667
100	3.594	40	4.07	40	4.07	32	3.9113333	37.33333333
300	10.782	60	12.21	92	12.21	68	11.734	73.33333333
1000	35.94	96	40.7	92	40.7	100	39.113333	96

(c)

Cb								
Dose (in µl)	Replicate 1		Replicate 2		Replicate 3		Plotted Data	
	cfu/ml	%mortality	cfu	%mortality	cfu	%mortality	cfu/ml-ave	%mortality-ave
1	0.03594	0	0.0407	0	0.0407	0	0.0391133	0
10	0.3594	4	0.407	8	0.407	6	0.3911333	6
30	1.0782	0	1.221	0	1.221	0	1.1734	0
100	3.594	0	4.07	4	4.07	2	3.9113333	2
300	10.782	0	12.21	0	12.21	0	11.734	0
1000	35.94	0	40.7	4	40.7	2	39.113333	2

(d)

pHT315 vector								
Dose (in µl)	Replicate 1		Replicate 2		Replicate 3		Plotted Data	
	cfu/ml	%mortality	cfu	%mortality	cfu	%mortality	cfu/ml-ave	%mortality-ave
1	0.05835	0	0.05865	0	0.05976	0	0.05892	0
10	0.5835	0	0.5865	0	0.5976	0	0.5892	0
30	1.7505	0	1.7595	0	1.7928	0	1.7676	0
100	5.835	0	5.865	0	5.976	0	5.892	0
300	17.505	0	17.595	0	17.928	0	17.676	0
1000	58.35	0	58.65	0	59.76	0	58.92	0

Table 3.4. Bioassay data against *Anopheles gambiae* used in Figure 3.4.A. (Cry operon and pHT315 cultures were grown for 24h at 30°C with shaking. Cbm and Cb cultures were grown in anaerobic jar for 18h at 30°C with no shaking. 25 third-instar larvae put in 200ml water in plastic cups were used for bioassay. %mortality was recorded after 24 hr.

Cry operon								
Dose (in µl)	Replicate 1		Replicate 2		Replicate 3		Plotted Data	
	cfu/ml	%mortality	cfu	%mortality	cfu	%mortality	cfu/ml-ave	%mortality-ave
1	0.05664	0	0.04553	0	0.05664	0	0.052937	0
10	0.5664	0	0.4553	0	0.5664	0	0.529367	0
30	1.6992	0	1.3659	0	1.6992	0	1.5881	0
100	5.664	0	4.553	0	5.664	0	5.293667	0
300	16.992	0	13.659	0	16.992	0	15.881	0
1000	56.64	0	45.53	0	56.64	0	52.93667	0
Cbm								
Dose (in µl)	Replicate 1		Replicate 2		Replicate 3		Plotted Data	
	cfu/ml	%mortality	cfu	%mortality	cfu	%mortality	cfu/ml-ave	%mortality-ave
0.01	0.000407	0	0.000407	0			0.000407	0
0.03	0.001221	13	0.001221	18			0.001221	15
0.1	0.00407	13	0.00407	19			0.00407	15
0.3	0.01221	21	0.01221	20			0.01221	20
1	0.0407	100	0.0407	100			0.0407	100
10	0.407	100	0.407	100			0.407	100
30	1.221	100	1.221	100			1.221	100
100	4.07	100	4.07	100			4.07	100
300	12.21	100	12.21	100			12.21	100
1000	40.7	100	40.7	100			40.7	100
Cb								
Dose (in µl)	Replicate 1		Replicate 2		Replicate 3		Plotted Data	
	cfu/ml	%mortality	cfu	%mortality	cfu	%mortality	cfu/ml-ave	%mortality-ave
1	0.03594	0	0.0407	0	0.0407	0	0.039113	0
10	0.3594	0	0.407	0	0.407	0	0.391133	0
30	1.0782	0	1.221	0	1.221	0	1.1734	0
100	3.594	0	4.07	0	4.07	0	3.911333	0
300	10.782	0	12.21	0	12.21	0	11.734	0
1000	35.94	0	40.7	0	40.7	0	39.11333	0
pHT315 vector								
Dose (in µl)	Replicate 1		Replicate 2		Replicate 3		Plotted Data	
	cfu/ml	%mortality	cfu	%mortality	cfu	%mortality	cfu/ml-ave	%mortality-ave
1	0.05835	0	0.05865	0	0.05976	0	0.05892	0
10	0.5835	0	0.5865	0	0.5976	0	0.5892	0
30	1.7505	0	1.7595	0	1.7928	0	1.7676	0
100	5.835	0	5.865	0	5.976	0	5.892	0
300	17.505	0	17.595	0	17.928	0	17.676	0
1000	58.35	0	58.65	0	59.76	0	58.92	0

Table 3.5. Data used for LC$_{50}$ calculation for toxicity of *Cb malaysia* and *Bacillus thuringiensis* expressing the *Cb malaysia* Cry operon to mosquito larvae shown in Table 3.6.

| Dose (in µl) | LC$_{50}$ in *Aedes* | | | | LC$_{50}$ in *Anopheles* | |
| | Cbm | | #9 | | Cbm | |
	cfu/ml	%mortality	cfu/ml	%mortality	cfu/ml	%mortality
0.01					0.000407	0.76
0.03					0.001221	7.24
0.1					0.00407	34.55
0.3					0.01221	71.59
1	0.039113	2.99	0.052937	7.18	0.0407	94.86
3	0.11734	7.84	0.15881	16.59	0.1221	99.99
10	0.391133	18.29	0.529367	33.31	0.407	100
30	1.1734	33.07	1.5881	52.42	1.221	100
100	3.911333	52.92	5.293667	72.57	4.07	100
300	11.734	70.53	15.881	86.25	12.21	100
1000	39.11333	85.33	52.93667	94.86	40.7	100

Table 3.6. Toxicity of *Cb malaysia* and *Bacillus thuringiensis* expressing the *Cb malaysia* Cry operon to mosquito larvae

	Aedes aegypti LC50, x108 cfu/ml (95% fiducial limits)	*Anopheles gambiae* LC50, x105 cfu/ml (95% fiducial limits)
Cb malaysia	3.30 (1.39 - 10.0)	6.39 (4.87 - 8.46)
Cry Operon	1.39 (0.70 - 2.74)	Non Toxic

Authors Note:

This chapter was published in Applied and Environmental Microbiology, 2014.

Qureshi, N., Chawla, S., Likitvivatanavong, S., Lee, H. L., and Gill, S.S. (2014). The cry toxin operon of *Clostridium bifermentans* subsp. *malaysia* is highly toxic to *Aedes* Larval Mosquitoes. *Appl Environ Microbiol*. **80**(18):5689-97.

Dr. Nadia Qureshi cloned the full length Cry Operon (pCryO) and *Cry16A* (pCry16) in

Bacillus expression vector pHT315 and showed its toxic activity and expression against

Aedes sp. is listed as the first co-author.

Dr. Supaporn Likitvivatanavong performed UPLC/MS/MS on 1- and 5-day cultures

proteins that were separated immediately by SDS – PAGE is listed as third author.

Chapter 4

Engineering *Culex* and *Anopheles* toxicity into *Clostridium bifermentans* subsp. *malaysia* Cry17Aa gene of the Cry operon.

Abstract

Improvement on mosquitocidal activity of *Clostridium bifermentans* subsp. *malaysia* Cry operon was planned by protein engineering of surface loop residues in domain II of Cry17Aa. The toxins were to be tested on 4 different species of important human disease vectors, *Aedes aegypti*, *Anopheles gambiae*, *Anopheles stephensi* and *Culex quinquefasciatus*. Here the strategy for achieving this objective is outlined.

Introduction

An anaerobic mosquitocidal bacterium *Clostridium bifermentans* subsp. *malaysia* (Cbm) has been studied by our lab to understand its mechanism of mosquitocidal activity. Cbm has been found to be highly toxic to larvae of several mosquito species (de Barjac et al., 1990; Thiery et al., 1992a). It is innocuous to mammals and goldfish (Thiery et al., 1992b). Cbm toxicity against *A. stephensi* is about 10 times higher, but its toxicity against *Culex pipiens* or *Aedes aegypti* is 10 times lower, than that of Bti (Thiery et al., 1992b). In Chapter 3, we showed that the larvicidal activity of Cbm against *Aedes aegypti* is selectively shown by Cry operon (Qureshi et al., 2014). Cry operon consists of unique combination of toxins, namely, Cry16A, Cry17A, Cbm17.1 and Cbm17.2, and that they act cooperatively.

Cry16A shows strong similarity within the toxic domain of *B. thuringiensis* delta-endotoxins, i.e., the amino-terminal part especially within blocks I to IV. The sequence of block V and the carboxy terminus of delta-endotoxins are absent from Cry16A (Barloy et al., 1996). Cry17A also shows similarities to Bt delta-endotoxins. Even though sequence similarities are unevenly distributed along the Cry17A protein, they are particularly common in the conserved Cry blocks, especially in blocks I and V. Block III is the least well conserved of the five (Barloy et al., 1998).

In Chapter 3, we showed that the Cry operon is highly toxic to *Aedes* and not toxic to *Anopheles* and *Culex*. We also showed by systematic deletion of genes from the original active Cry operon, that all genes within the Cry operon are required for larvicidal activity (Figure 4.9). In this study we show that the toxin complex lost its toxicity when

a region in Cry17A was deleted using MscI-ApaI restriction digest to create the CryO/Δ17. In this chapter, we wanted to investigate the importance of this region by swapping the deleted region with that of another pore forming Cry toxin. We also tried to introducing short variable sequences of the loop regions from one toxin to another might provide a rational design approach of engineering the activity of Cry toxins in *Anopheles* and *Aedes*. Our approach was to generate three dimensional structures of mosquitocidal toxins by homology modeling to study the deleted and domain II loop regions of Cry17A and Cry4Ba in search of sites involved in mosquitocidal activities. The exploration would start by exchanging loop residues from Cry4Ba to Cry17A. The roles of the loop exchanges were then to be tested on 4 different species of important human disease vectors, *Aedes aegypti* (dengue, yellow fever), *Anopheles gambiae* (malaria), *Anopheles stephensi* (malaria) and *Culex quinquefasciatus* (West Nile virus).

Material and Methods

Bacterial strains and culture conditions

C. bifermentans (ATCC) was used as the wild type reference strain and *C. bifermentans* serovar *malaysia* (*Cb malaysia*) was from the collection of the Institute for Medical Research, Malaysia (Lee and Seleena, 1990b). These strains were used for bioassays and for isolation of gDNA for cloning. *Escherichia coli* strainS DH10 beta electrocompetent cells (New England Biolabs, Ipswich, MA) were used for cloning and the NEB transformation protocol was used. *B. thuringiensis* subsp. *israelensis* 4Q7

146

(Bacillus Stock Center, Ohio State University, Columbus, Ohio) was used for expression of toxin genes using a published protocol (Macaluso and Mettus, 1991). Clostridial strains were streaked for isolation on TYG agar and liquid cultures were grown in TYG medium at 30°C under anaerobic conditions using BD GasPakEZ (Beckton-Dickinson Microbiology, Cockeysville, MD). The liquid cultures for *Clostridium* strains were grown anaerobically at 30°C, while the *Bacillus* cultures were grown at 30°C with shaking. Nutrient broth media was used for *B. thuringiensis* and Luria broth media was used for *E. coli*. Ampicillin (100μg/ml) and erythromycin (25μg/ml) were added when required.

General DNA Techniques

All DNA manipulations were performed according to standard protocols. Primers used in this study were created using the sequences from existing databases. Polymerase chain reactions (PCR) were performed in an automated thermacycler (C 1000 Touch ™, Bio Rad). Choice-TaqTM mastermix DNA polymerase (Denville Scientific, Metuchen, NJ) was used for all PCR reactions for products below 2kb. Advantage 2x polymerase (Clontech, Mountain View, CA) or Phusion polymerase (NEB) was used for all PCR reactions for products above 2kb. PCR products were separated in 1% agarose gels and subsequently cut and purified using Wizard SV Gel and PCR purification kits (Promega, Madison, WI). Sequencing of purified DNA products was performed by the Genomics Core facility at the University of California, Riverside.

Homology modeling of Cry operon toxins from *C. bifermentans* subsp. *malaysia*

Two programs were used for modeling purposes. First, an internet-based CLUSTAL W version (McWilliam et al., 2013) available at (http://www.ebi.ac.uk/Tools/msa/clustalo/). Second, SWISS-MODEL (Arnold et al., 2006) available at (http://www.expasy.org/swissmod/SWISS-MODEL.html). Multiple sequence alignments were performed using CLUSTALW. Models were constructed using the "Optimize (project) mode" as well as the "Automated (project) mode" in SWISS-MODEL, and viewed with Swiss-Pdb Viewer. The template file used was chosen from either the known tertiary structures or from the models that were successfully constructed with SWISS-MODEL. Models of the four Cry operon proteins were constructed with this technique. Also, a model of Cry17Aa was constructed using the model structure of Cry4Ba (PDB: 1W99) as a template. The two sequences were aligned using CLUSTAL W and the residues of Cry17Aa were substituted with loop residues of Cry4Ba by site-directed mutagenesis.

Cry17A domain I Swap by Cry4Ba domain I

624bp fragment corresponding α2b to α8b domain from Cry4Ba gene with correct orientation with MscI and ApaI overhangs was cloned in a pCR-2.1 vector using primers 1 and 2 (Table 4.1) to create TA-Cry4Ba. The clones were sequenced completely and the correct clone was selected. pCryO clone and TA-Cry4Ba was digested with MscI and ApaI sequentially. The digest was run on 1% agarose gel and the correct sized fragments were cut and purified using the Wizard SV Gel and PCR purification kits (Promega,

Madison, WI). The purified fragments were ligated and transformed in XL-10 cells

(NEB). Correct clone was to be selected by using *DraIII* digest and sequencing using the

PCR primers. The selected clone was then to be transform in Bt-4Q7 and bioassayed.

Mutating Cry17Aa by site-directed mutagenesis

A 2kb region surrounding the 636bp MscI-ApaI deleted region was cloned in a

pCR-2.1 vector using primers 3 and 4 (Table 4.1) with specific restriction sites BglII and

StuI. The clones were sequenced completely and the correct clone was selected for

mutagenesis protocol. Site-directed mutagenesis was performed using the GeneArt®

Site-Directed Mutagenesis System (Invitrogen) by following manufactures recommended

protocol. DNA templates were purified using a SV Wizard plasmid purification kit

(Promega). Adavantage 2 Polymerase (Clontech) was used for PCR. The mutagenized

clones were transformed in DH10 beta electrocompetent cells. Plasmid DNA was

extracted using the SV Wizard plasmid purification kit (Promega) and sent for

sequencing. The clones showing the correct mutagenized sequences were to be selected

and digested with BglII and StuI. This mutagenized fragment was then to be replaced in

the original pCryO clone from Chapter 3. The mutagenized pCryO clone was to be

transformed in DH10 beta cell. The clones were to be screened by colony PCR and

sequencing for the presence of the mutagenized plasmid. The correct mutagenized pCryO

would be used to transformed Bt 4Q7 cells.

Bioassays

Total cultures of Cbm, Cb and recombinant *B. thuringiensis* (rBT) strains pCryO-Cry4ba and CryO/Δ17 and loop-mutant Cry17A strains were assayed against *Ae. aegypti, An. gambiae* , and *C. quinquefasciatus* larvae. Bioassays were performed at 24° C using 25 third instar larvae in 200 ml water. Live bacterial cultures were used in bioassays. Larval mortality was determined by counting the number of larvae still alive at 24 or 48h. Bioassays were repeated 2-3 times and LC_{50} concentrations were determined by probit analysis (USDA) and plotted using the Origin program (Origin Lab, Northampton, MA).

Antibodies

Cry16A antibody was designed as previously published in Qureshi et al., 2014. Affinity purified Cry17A peptide (NNKKIEQNKIVEYNC) polyclonal antibody was commercially synthesized (GenScript USA, Piscataway, NJ).

Polyacylamide gel electrophoresis (PAGE) and immunoblot

C. bifermantans, *Cb malaysia*, rBT and 4Q7 strains were grown in parallel for 18h. Aliquots, 100ul, of the cultures were mixed with of 2X sample buffer (Bio-Rad, Hercules, CA), and the samples separated by SDS-polyacrylamide (10-14%) gel electrophoresis. Native polyacrylamide (10%) gel electrophoresis was performed using 1X native PAGE buffer (3g Tris, 14.5g glycine). In parallel the separated proteins were transferred onto PVDF membranes (Millipore, Billerica, MA). The membranes were treated with blocking solution (1X PBS, 5% skim milk and 0.1% Tween-20) for 1 h at

room temperature and then washed with PBST (1XPBS and 0.1% Tween-20). The

blocked membrane was incubated overnight (at 4°C) with primary anti-Cry16A or

Cry17A antibody at 1:1000 dilution. A secondary antibody, ECL™ anti-rabbit IgG-

horseradish peroxidase-linked whole antibody (GE Healthcare, Anaheim, CA) was used

at a dilution of 1:5000. Immunoreactive bands were visualized using the ECL western

blotting kit (GE Healthcare) and exposed to an X-ray film.

Results

Homology Modelling of Cry Operon proteins

I attempted to create models of the four Cry Operon proteins, Cry16A, Cry17A,

Cbm 17.1 and Cbm 17.2. SWISS-MODEL gives all the models a GQME score. GMQE

(Global Model Quality Estimation) is a quality estimation which combines properties

from the target-template alignment. The resulting GMQE score is expressed as a number

between zero and one, reflecting the expected accuracy of a model built with that

alignment and template. Higher numbers indicate higher reliability. The Cry models

received a score of 0.59 showing better homology than the Cbm 17 toxins models that got

a score of about 0.39. Figure 4.1 depicts the analysis of putative model structure of 617aa

long Cry16A protein based on Cry8Ea1with a score of 0.58 and Figure 4.2 depicts the

analysis of putative model structure of 618aa long Cry17A protein based on Cry3Aa.

Figures 4.3and 4.4 shows the analysis of putative model structures of 152aa long

Cbm17.1 protein and 153aa long Cbm17.2 respectively, both based on Cry34Ab1.

Homology modeling was performed to obtain model structure of Cry17A using Cry4Ba

as the template (Figure. 4.5). This model shows that Domain I of Cry17A and Cry4Ba are very similar, as indicated by the blue color. Cry4Ba Domain I is a 290 aa, seven helix bundle, putative pore forming domain. Clustal W analysis of the Cry17A and Cry4Ba protein sequences (Figure 4.6) showed that 636 bp Fragment deletion of Cry17A gene is a deletion of the whole domain from α2b to α8b.

Expression and Larvicidal Activity of CryO/Δ17 deletion construct

CryO/Δ17 deletion construct was made in our lab by restricting pCryO with MscI and ApaI followed by relegation. This restriction deleted the Cry17A from pCryO. Protein expression from this transformation, rBT pCryO/Δ17, was analyzed by Coomassie Blue staining and western blotting (Figure 4.10). Cultures of pCryO and pCryO/Δ17 were grown for 18h. Both the proteins were expressed. A strong band was observed around 70kDa for the pCryO with anti-Cry16 antibody (Figure 4.10A), whereas pCryO/Δ17 showed degradation at 18h. Whole cell preparations from Cbm, Cb, pCryO/Δ17 and rBT transformants containing pHT315 (control) and pCryO that expressed the entire Cry operon were assayed for larvicidal activity using third instar *Ae. aegypti*. The rBT pHT315 control samples showed no larval toxicity against all three mosquito species. pCryO was toxic to *Aedes aegypti* as shown before but pCryO/Δ17 showed no toxicity to the same.

In Chapter 3, we hypothesized that a complex may be needed for toxicity and looked to determine whether pCryO/Δ17 formed that complex. Indeed analysis by native PAGE, pCryO/Δ17 showed a complex of about 170kDa (Figure 4.10B). In contrast, as

152

shown in Chapter 3, rBT containing pCry16, pCry16/17 or pCryO/Δ17 did not show the formation of a significant amount of the complex. Therefore, this result can be summarized as shown in Figure 4.9.

Discussion

The ability to improve mosquitocidal activity by rational design opens new avenues for genetic manipulation for optimization of biological control of mosquitoes. In Chapter 3, we hypothesized that Cry Operon proteins form a complex and this Large Toxin Complex (LC) is necessary for toxicity. To analyze the key genes involved in toxicity, individual Cry operon gene constructs were made and expressed. But individual toxins are by themselves not toxic. So, a few more constructs were made by deleting one or more genes, to help identify the role of specific genes involved in toxicity (summarized in Figure 4.9). pCryO/ΔCry17A, showing *Cry17A* gene deletion from Cry operon, was made by Stefani Andrews, a graduate student in our lab. In the process of creating the pCryO/ΔCry17A construct, she deleted a 636bp MscI-ApaI restriction fragment from the middle of the *Cry17A* gene. The expression of pCryO/ΔCry17A seemed unstable and the protein showed degradation at conditions in which pCryO was stably expressed. This construct although not toxic to *Aedes*, showed the formation of 170kda complex on the native gel. This result compelled us to ask the question, what is the function of this 636 bp fragment, since this minimum loss of a fragment makes the whole Cry complex not toxic, yet show complex formation.

By protein modelling, we found that Cry17A shows some homology to a pore forming Cry4Ba toxin. We then obtained the model of Cry17A using Cry4Ba as a template (Figure 4.5) using SWISS-MODEL "Automated Project Mode". This model shows that Domain I of Cry17A and Cry4Ba are very similar. Clustal W analysis of the Cry17A and Cry4Ba protein sequences showed that 636 bp Fragment deletion of *Cry17A* gene is a deletion of the whole domain I from α2b to α8b (Figure 4.6). Therefore, I propose two experiments to answer this question. First, by investigating the possibility that exchanging pore forming helices from one Cry toxin to another pore forming Cry toxin that shows toxicity to mosquitoes (like Cry4Ba) might also transfer the toxicity determinants as well. For this experiment, I chose to swap the deleted portion of Cry17A with that of mosquitocidal Cry4Ba toxin as it is also a pore forming toxin. Second, by investigating the possibility of exchanging loop residues from Cry17A toxin to Cry4Ba toxin might also transfer the toxicity determinants as well. The objective of the loop exchange experiment was to introduce *Anopheles* and *Culex* toxicity into Cry Operon by matching as closely as possible to the apical loop sequences and the length of the loops in Cry4Ba and Cry4Aa respectively, by mutagenesis.

Cry4Ba Domain I consists of the first 290 amino acids that form A seven helix bundle. It has been ascertained that this region is a putative pore forming domain. Therefore, the objective of swapping the 213 amino acid MscI-ApaI fragment of CryO/ΔCry17A with 230 amino acid fragment from Cry4Ba that covers α2b to α8b, was to re-introduce the deleted pore forming region to the CryO/ΔCry17A clone. I hypothesize that the exchange of the pore forming domain would result in an equal if not

increased activity of the mutant toxin that could be detected through similar or increased *Anopheles* mortality when compared to the native pCryO.

The second domain of the 3-domain Cry toxins is composed of anti-parallel β-sheets and is known to bind carbohydrates through the exposed loops at the apex of its β-prism fold (Li et al., 1991). The three surface exposed loops of domain II also show similarities to immunoglobin antigen-binding sites and were proposed to be involved in receptor binding and hence species specificity. Site-directed mutagenesis and segment swapping experiments have supported this model (Abdullah et al., 2003). The effects of mutations in sequences encoding domain II loop regions in Cry1Aa, Cry1Ab, Cry1Ac, and Cry3Aa (Schnepf et al., 1998) have shown that the mutations may have a negative or positive effect on binding and toxicity. Therefore, there is a possibility that exchanging loop residues from one Cry toxin to another Cry toxin might also transfer the toxicity determinants as well. Cry17A and Cry4Ba, a pore forming toxin, were chosen based on their differences in specificity towards different species of mosquitoes. Cry4Aa has been reported to be moderately toxic to *Aedes*, *Anopheles*, and *Culex* larvae, while Cry4Ba has been reported to be very toxic to *Aedes* and *Anopheles* larvae but not toxic at all to *Culex* larvae (Angsuthanasombat et al., 1991, 1992). The Cbm Cry operon complex has been shown previously to be toxic to *Aedes* but not toxic to *Anopheles* and *Culex* (Qureshi et al., 2014). Putative models of both Cry17A and Cry4Ba toxins obtained through homology modeling were examined. From the models, putative surface exposed loop residues in domain II were determined and it was apparent that not only the sequences were different, the length of the loops were different as well between the two toxins.

Cry4A and Cry4Ba were chosen based on the near homology of their sequences, and their differences in specificity towards different species of mosquitoes as compared to Cry17A. Cry4Ba is highly toxic to *Anopheles* mosquitoes and Cry4Aa is toxic to *Culex* whereas Cry operon is not (Qureshi et al., 2104). Also, my lab has stable, over-expressing clones in Bt 4Q7 cells for this toxin.

Two loop regions were identified with loop 3 of Cry4Ba slightly larger than the corresponding loops of Cry17A. The objective of the loop exchange experiment was to introduce *Anopheles* toxicity into Cry17A by matching as closely as possible to the apical loop sequences and the length of the loops in Cry4Ba. Ibrahim et al., 2013 showed that 447- YIKTDVIDYN -458 residues in Cry4Ba (Figure 4.8) play an important role in toxicity to *Anopheles*. Upon multiple sequence alignment of Cry16A, Cry17A and Cry4Ba, it was observed that Cry17A shows the same motif 414- YKMHHYGDNY – 424. The approach was to mutate Tyr-420 to Valine (UAU to GUU).

In a similar experiment, the loop 3 sequence of Cry17A would be exchanged with the loop 3 sequence of Cry4Aa gene, to test the hypothesis that *Culex* activity from Cry4Aa could be transferred to Cry17A by exchanging the loop regions of domain II. Abdullah et al., 2003 showed that 511-PATYK-516 residues in Cry4Aa play an important role in toxicity to *Culex* by changing DYN in 447-YIKTDVIDYN - 458 in Cry4Ba to PATY. Upon multiple sequence alignment of Cry16A, Cry17A and Cry4Aa, it was observed that Cry17A shows the same motif 414- YKMHHYGDNY – 424. The approach was to mutate D→P (AGC→ CCC), followed by inserting A (GCC), followed by changing N→T (AAC→ACC).

156

Several functions are assigned to the three-domain structure of Cry toxins. Domain I functions in the formation and function of ion channel (Schnepf et al., 1998). Domains II and III have been linked to specific binding to receptors (Aronson et al., 2001; Schnepf et al., 1998). Domain III has been implicated in modulation of ion channel activity (MacKinnon et al., 1989, 1990). Since loop regions in domain II are correlated to specificity in some cases, and mutations in the sequence of loop regions affected toxicity, there is a possibility that exchanging loop residues from Cry17A to Cry4Ba toxin might also transfer the *Anopheles* toxicity determinants as well. As it proved inefficient to do site-directed mutagenesis on large 12kb plasmid, I decided to clone the targeted region in a smaller pCR2.1-TOPO vector to obtain a 5kb plasmid for mutagenesis. At first, I decided on exchanging 3 amino acid residues at each mutagenesis step instead of single amino acid residue mutations to maximize the effect of the mutations. Increasing the number of mutations per mutagenesis reaction would decrease the efficiency of the reaction and limiting the number of mutations to three residues per mutagenesis reaction would alleviate this problem. Since the first approach failed, the second approach would be to do single amino mutations at each step. The approach taken to enhance the toxin was by domain II loop-exchanges with another toxin with known toxicity. Unfortunately, I have not been able to get a mutant clone either for Loop3-*Anopheles* change or Loop3-*Culex* change mutagenesis experiments.

I hypothesize that the Cry17A loop 3 exchanges to Y---V—Y and PATYK residues would cause significant increase in toxicity against *Anopheles* and *Culex* respectively, while still maintaining activity against *Aedes* mosquitoes. The increase in

toxicity could be related to the ability of the mutant proteins to bind to *Anopheles* and

Culex BBMV due to the presence of the "new" motif residues. We should see that the

mutant toxins are able to bind to *Anopheles* and *Culex* BBMV whereas the native CryO is

not toxic.

References

Abdullah, M. A., 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. *AEM*, **69**(9): 5343–5353

Angsuthanasombat, C., Crickmore, N. and Ellar, D. J. (1991). Cytotoxicity of a cloned *Bacillus thuringiensis* subsp, *israelensis* CryIVB toxin to an *Aedes aegypti* cell line. *FEMS Microbiol Lett.***83**: 273-276.

Angsuthanasombat, C., Crickmore, N. and Ellar, D. J. (1992). Comparison of *Bacillus thuringiensis* subsp, *israelensis* CrylVA and CryIVB cloned toxins reveals synergism in vivo. *FEMS Microbiol Lett.***94**: 63-68

Arantes, O. and Lereclus, D. (1991). Construction of cloning vectors for *Bacillus thuringiensis*. *Gene*. **108**: 115-119.

Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics.***22**:195-201.

Aronson, A. I., and Shai, Y. (2001). Why *Bacillus thuringiensis* insecticidal toxins are so effective: unique features of their mode of action. *FEMS Microbiol Lett.***195**:1-8.

Barloy, F., Delecluse, A., Nicolas, L., and Lecadet, M.M. (1996). Cloning and expression of the first anaerobic toxin gene from *Clostridium bifermentans* subsp. *malaysia*, encoding a new mosquitocidal protein with homologies to *Bacillus thuringiensis* delta-endotoxins. *J Bacteriol*. **78**: 3099-3105.

Barloy, F., Lecadet, M.M., and Delecluse, A. (1998a). Distribution of clostridial cry-like genes among *Bacillus thuringiensis* and *Clostridium* strains. *Curr Microbiol*. **36**: 232-237.

Barloy, F., Lecadet, M.M., and Delecluse, A. (1998b) Cloning and sequencing of three new putative toxin genes from *Clostridium bifermentans* CH18. *Gene*. **211**: 293-299.

Berne S, Lah L, Sepcic K. (2009). Aegerolysins: structure, function, and putative biological role. *Protein Sci*. **18**:694-706.

Chang, C., Yu, Y.M., Dai, S.M., Law, S.K., and Gill, S.S. (1993). High-level cryIVD and cytA gene expression in *Bacillus thuringiensis* does not require the 20-kilodalton protein, and the coexpressed gene products are synergistic in their toxicity to mosquitoes. *Appl Environ Microbiol*. **59**: 815-821.

Charles, J.-F., and Nielsen-LeRoux, C. (2000). Mosquitocidal bacterial toxins: Diversity, mode of action and resistance phenomena. *Mem Institut Oswaldo Cruz.* **95**: 201-206.

Charles, J.F., Nicolas, L., Sebald, M., and de Barjac, H. (1990). *Clostridium bifermentans* serovar *malaysia*: sporulation, biogenesis of inclusion bodies and larvicidal effect on mosquito. *Research Microbiol.* **141**: 721-733.

Crickmore, N., Zeigler, D.R., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J., Bravo, A. and Dean, D.H. (2014). *Bacillus thuringiensis* toxin nomenclature. http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/index.html

de Barjac, H., Sebald, M., Charles, J.F., Cheong, W.H., and Lee, H.L. (1990). *Clostridium bifermentans* serovar *malaysia*, a new anaerobic bacterium pathogen to mosquito and blackfly larvae. *C R Acad Sci III*. **310**: 383-387.

de Maagd, R. A., Bravo, A. and Crickmore, N.(2001). How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. *Trends Genet* **17**:193-9.

Galitsky, N., Cody,v., Wojtczak, A. , Ghosh, D., Luft, J. R., Pangborn, W. and English., L. (2001). Structure of the insecticidal bacterial delto-endotoxin Cry3Bb1 of *Bacillus thuringiensis*. *Acta Crystallogr D Biol Crystallogr* **57**:1101-9.

Georghiou, G.P., and Wirth, M.C. (1997). Influence of exposure to single versus multiple toxins of *Bacillus thuringiensis* subsp. *israelensis* on development of resistance in the mosquito *Culex quinquefasciatus* (Diptera: Culicidae). *Appl Environ Microbiol*. **63**: 1095-1101.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods.* **6**: 343-345.

Ibrahim, M. A., Griko, N. B. and Bulla, Jr. L.A. (2013). The Cry4B toxin of *Bacillus thuringiensis* subsp. *israelensis* kills Permethrin-resistant *Anopheles gambiae*, the principal vector of malaria. *Exp Biol Med (Maywood)*. **238**(4):350-9.

Juarez-Perez, V., and Delecluse, A. (2001). The Cry toxins and the putative hemolysins of *Clostridium bifermentans* serovar *malaysia* are not involved in mosquitocidal activity. *J Invertebr Pathol*. **78**: 57-58.

Lederberg, J., Shope R.E., Oakes S.C. (1992). Emerging infections: microbial threats to health in the United States. *Washington: National Academy Press*

Lee, H.L., and Seleena, P. (1990a). Isolation of indigenous larvicidal microbial control agents of mosquitos: the Malaysian experience. *Southeast Asian J Trop Med Public Health*. **21**: 281- 287.

Lee, H.L., and Seleena, P. (1990b). Isolation and evaluation of larvicidal *Clostridium bifermantans* against mosquitoes of public health importance. *Trop Biomed*. **7**: 103-106.

Li, J., Koni, P. A. and Ellar, D. J. (1996). Structure of the mosquitocidal delta-endotoxin CytB from *Bacillus thuringiensis* sp. *kyushuensis* and implications for membrane pore formation. *J Mol Biol* **257**: 129-152.

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**: 815-821.

Macaluso, A., and Mettus, A.M. (1991). Efficient transformation of Bacillus thuringiensis requires nonmethylated plasmid DNA. *J Bacteriol.* **173**: 1353-1356.

MacKinnon, R., and C. Miller. (1989). Mutant potassium channels with altered binding of charybdotoxin, a pore-blocking peptide inhibitor. *Science*. **245**:1382-1385.
MacKinnon, R., and G. Yellen. (1990). Mutations affecting TEA blockade and ion permeation in voltage-activated K+ channels. *Science*. **250**:276-279.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. and Lopez, R. (2013). Analysis Tool Web Services from the EMBL-EBI. *Nucleic acids research(Web Server issue)* **41**:W597-600

Murray, C.J., Rosenfeld, L.C., Lim, S.S., Andrews, K.G., Foreman, K.J., Haring, D. (2012). Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet.* **379**: 413- 431.

Murray NE, Quam MB, Wilder-Smith A. (2013). Epidemiology of dengue: past, present and future prospects. Clin Epidemiol **5:**299-309.

Najera, J.A. (2001). Malaria control: achievements, problems and strategies. *Parassitologia*. **43**: 346 1-89.

Nicolas, L., Hamon, S., Frachon, E., Sebald, M., and Barjac, H. (1990). Partial inactivation of the mosquitocidal activity of *Clostridium bifermentans* serovar *malaysia* by extracellular proteinases. *Applied Microbiol. Biotechnol.* **34**: 36-41.

Nicolas, L., Charles, J.F., and de Barjac, H. (1993). *Clostridium bifermentans* serovar *malaysia*: characterization of putative mosquito larvicidal proteins. *FEMS Microbiol Lett.* **113**: 23-28.

Nielsen-Leroux, C., Charles, J.F., Thiery, I., and Georghiou, G.P. (1995). Resistance in a laboratory population of *Culex quinquefasciatus* (Diptera: Culicidae) to *Bacillus sphaericus* binary toxin is due to a change in the receptor on midgut brush-border membranes. *Eur J Biochem.* **228**: 206-210.

Rao, D.R., Mani, T.R., Rajendran, R., Joseph, A.S., Gajanana, A., and Reuben, R. (1995). Development of a high level of resistance to *Bacillus sphaericus* in a field population of *Culex quinquefasciatus* from Kochi, India. *J Am Mosq Control Assoc.* **11**: 1-5.

Roberts, D.R., Manguin, S., and Mouchet, J. (2000). DDT house spraying and re-emerging malaria. *Lancet.* **356**: 330-332.

Sankaranarayanan, R., K. Sekar, R. Banerjee, V. Sharma, A. Surolia, and M. Vijayan. (1996). A novel mode of carbohydrate recognition in jacalin, a *Moraceae* plant lectin with a β-prism fold. *Nat. Struct. Biol.* **3**:596-603.

Schnepf, E., Crickmore, N., VanRie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D. R. and Dean, D. H. (1998). *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* **62:**775-806.

Thiery, I., Hamon, S., Gaven, B., and De Barjac, H. (1992) Host range of *Clostridium bifermentans* serovar *malaysia*, a mosquitocidal anaerobic bacterium. *J Am Mosq Control Assoc.* **8**: 272-277.

World Health Organization. (2004). World Health Report. WHO, Geneva.

Wu, D., and Federici, B.A. (1993) A 20-kilodalton protein preserves cell viability and promote CytA crystal formation during sporulation in *Bacillus thuringiensis*. *J Bacteriol.* **175**: 5276- 5280.

Yiallouros, M., Storch, V., Thiery, I., and Becker, N. (1994). Efficacy of *Clostridium bifermentans* serovar *malaysia* on target and nontarget organisms. *J Am Mosq Control Assoc.* **10**: 51-55.
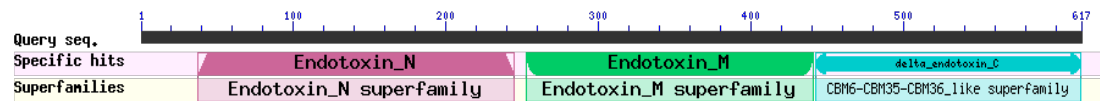
Cry16A

Nucleotide sequence: 1865 bp.

Protein Sequence: 617 a.a.

Computed pI: 5.06

blastP: 30% identical to Cry19Aa (accession no. O32307.1)

     31% identical to Cry27Aa (accession no. Q9S597.1)

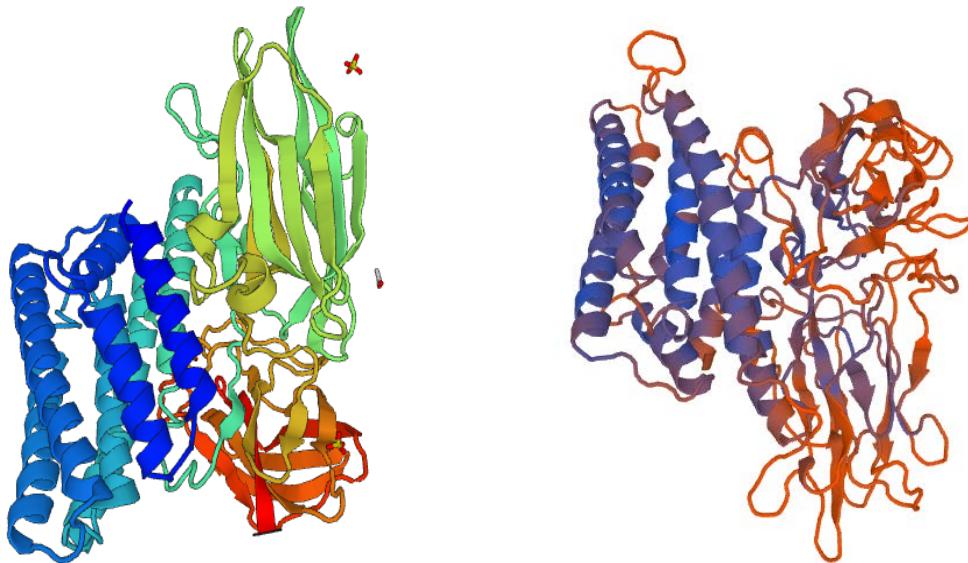**Cry Toxin Group and Toxicity**: Cry-Three domain toxins; Dipterans (Flies)

SWISS-Model: Insecticidal Delta-Endotoxin Cry8Ea1

      Template name: 3eb7.1.A

      Sequence Coverage: 89%

      Sequence Similarity: 34%

      GMQE score: 0.58

Insecticidal Delta-Endotoxin Cry8Ea1

Figure 4.1: Homology Modelling of Cry16A. Analysis of Cry16A gene and protein sequence with a putative model of Cry16A based on Insecticidal Delta – Endotoxin Cry8Ea1 (Guo, S. et al., 2009; PDB: 3eb7.1.A) using SWISS-Model "Automated Project Mode".

GMQE (Global Model Quality Estimation) is a quality estimation which combines properties from the target-template alignment. The resulting GMQE score is expressed as a number between zero and one, reflecting the expected accuracy of a model built with that alignment and template. Higher numbers indicate higher reliability.

For Cry16A model, Red indicates highly dissimilar. Blue indicates highly similar. In between shades: sliding scale

<u>Cry17A</u>

Gene sequence: 1857 bp
Protein Sequence: 618 aa
Computed pI: 5.05
blastP: 33% identical to Cry27Ab1 (accession no. AGV55018.1)
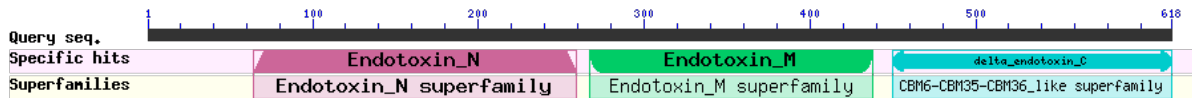
Cry Toxin Group and Toxicity: Cry-Three domain toxins; Dipterans (Flies)

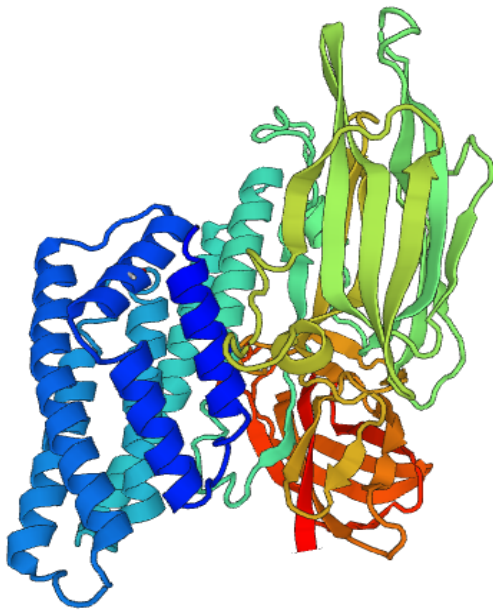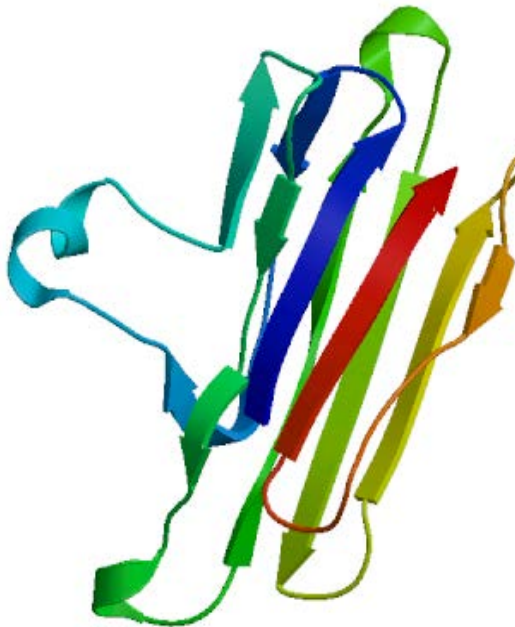**SWISS-Model:**   <u>Pesticidal crystal protein cry3Aa</u>
                      Template name: 4qx0.1.A
                      Sequence Coverage: 88%
                      Sequence Similarity: 34%
                      GMQE score: 0.59

Pesticidal crystal protein Cry3Aa

Cry17A model based on Cry3Aa

Figure 4.2: Homology Modelling of Cry17A. Analysis of Cry17A gene and protein sequence with a putative model of Cry17A based on Pesticidal crystal protein Cry3Aa (Sawaya et al., 2014; PDB: 4qx0.1.A) using SWISS-Model "Automated Project Mode".

GMQE (Global Model Quality Estimation) is a quality estimation which combines properties from the target-template alignment. The resulting GMQE score is expressed as a number between zero and one, reflecting the expected accuracy of a model built with that alignment and template. Higher numbers indicate higher reliability.

For Cry17A model, Red indicates highly dissimilar. Blue indicates highly similar. In between shades: sliding scale

Cbm17.1

Gene sequence: 462 bp

Protein Sequence: 153 aa

Computed pI: 5.31

blastP: 78% identical to Cbm17.2

              45% identical to *Pseudomonas aeruginosa* hemolysins

              (Accession nos. WP_014602345.1 & WP_003083767.1)

Cry Toxin Group and Toxicity: Cry non-related to other toxins;

                                        Coleopterans (beetles)

SWISS-Model: 13.6 kDa insecticidal crystal protein Cry34Ab1

              Template name: 4jox.1.A
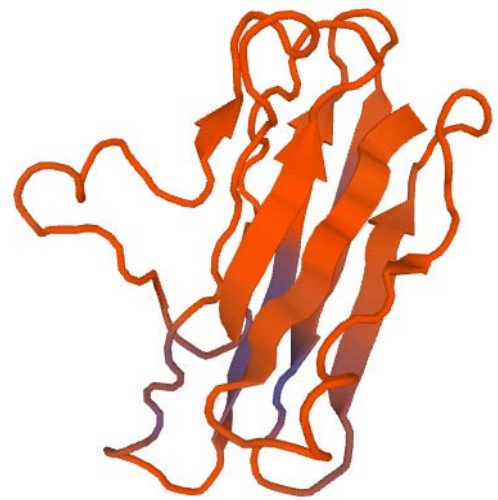
              Sequence Coverage: 68%

              Sequence Similarity: 33%

              GMQE score: 0.39

Insecticidal crystal protein Cry34Ab1

Cbm17.1 model based on Cry34Ab1

Figure 4.3: Homology Modelling of Cbm17.1. Analysis of Cbm17.1 gene and protein sequence with a putative model of Cbm17.1 based on 13.6kDa insecticidal crystal protein Cry34Ab1 (Kelkar. et al., Unpublished; PDB: 4jox.1.A) using SWISS-Model "Automated Project Mode".

GMQE (Global Model Quality Estimation) is a quality estimation which combines properties from the target-template alignment. The resulting GMQE score is expressed as a number between zero and one, reflecting the expected accuracy of a model built with that alignment and template. Higher numbers indicate higher reliability.

For Cbm17.1 model, Red indicates highly dissimilar. Blue indicates highly similar. In between shades: sliding scale

Cbm17.2

Gene sequence: 459 bp

Protein Sequence: 152 aa

Computed pI: 5.32

blastP: 78% identical to Cbm17.1

45% identical to *Pseudomonas aeruginosa* BWH057

hemolysins (accession no. EZO68291.1)

Cry Toxin Group and Toxicity: Cry-Bin like; Coleopterans (beetles)

SWISS-Model: 13.6 kDa insecticidal crystal protein Cry34Ab1
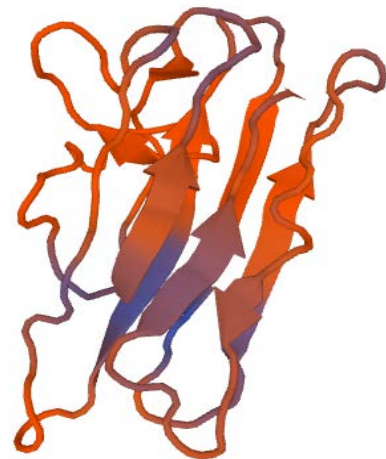
Template name: 4jox.1.A

Sequence Coverage: 68%

Sequence Similarity: 33%

GMQE score: 0.39

Insecticidal crystal protein Cry34Ab1   Cbm17.2 model based on Cry34Ab1

Figure 4.4: Homology Modelling of Cbm17.2. Analysis of Cbm17.2 gene and protein sequence with a putative model of Cbm17.2 based on 13.6kDa insecticidal crystal protein Cry34Ab1 (Kelkar. et al., Unpublished; PDB: 4jox.1.A) using SWISS-Model "Automated Project Mode".

GMQE (Global Model Quality Estimation) is a quality estimation which combines properties from the target-template alignment. The resulting GMQE score is expressed as a number between zero and one, reflecting the expected accuracy of a model built with that alignment and template. Higher numbers indicate higher reliability.

For Cbm17.2 model, Red indicates highly dissimilar. Blue indicates highly similar. In between shades: sliding scale
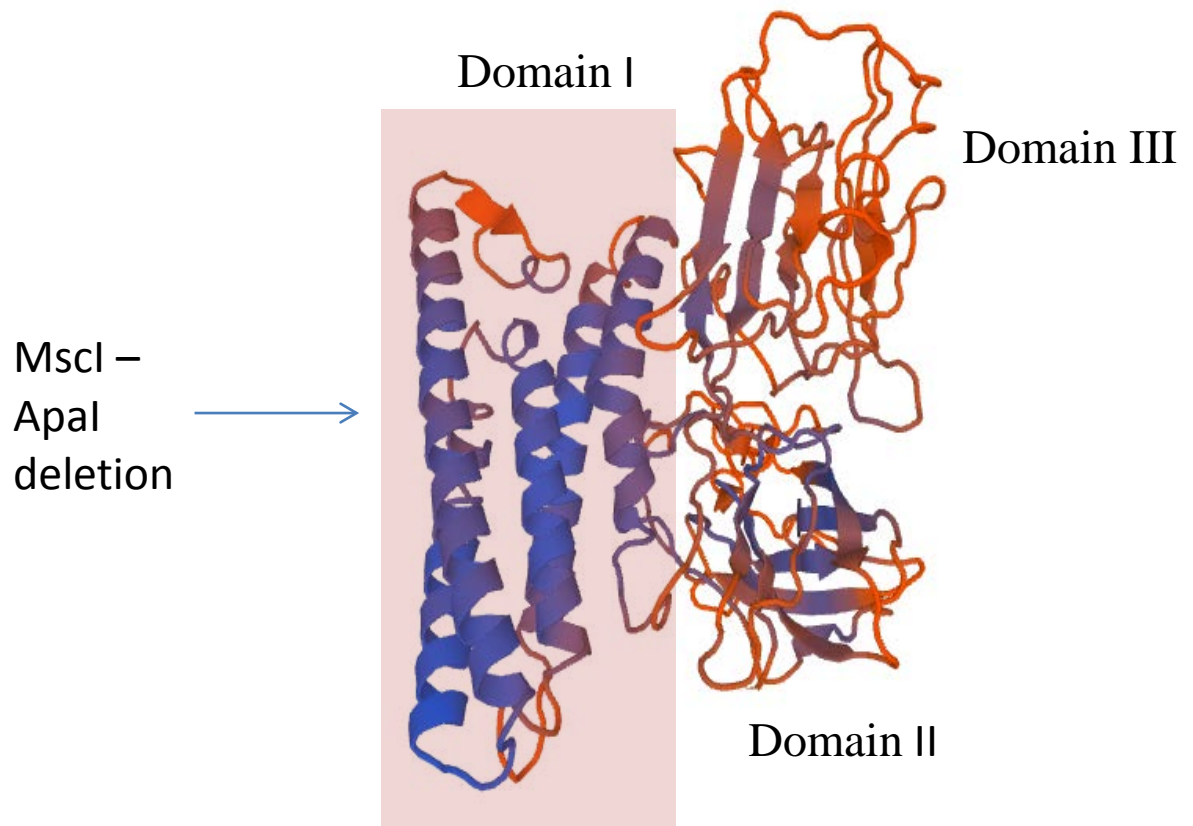
Domain I

Domain III

MscI –
ApaI
deletion

Domain II

Figure 4.5: Homology Modelling of Cry17A for site directed mutagenesis study. A
putative model of Cry17A was created based on Cry4Ba (Boonserm et al., 2005; PDB:
1W99.A) using SWISS-Model "Optimized Project Mode". Shaded region depicts the
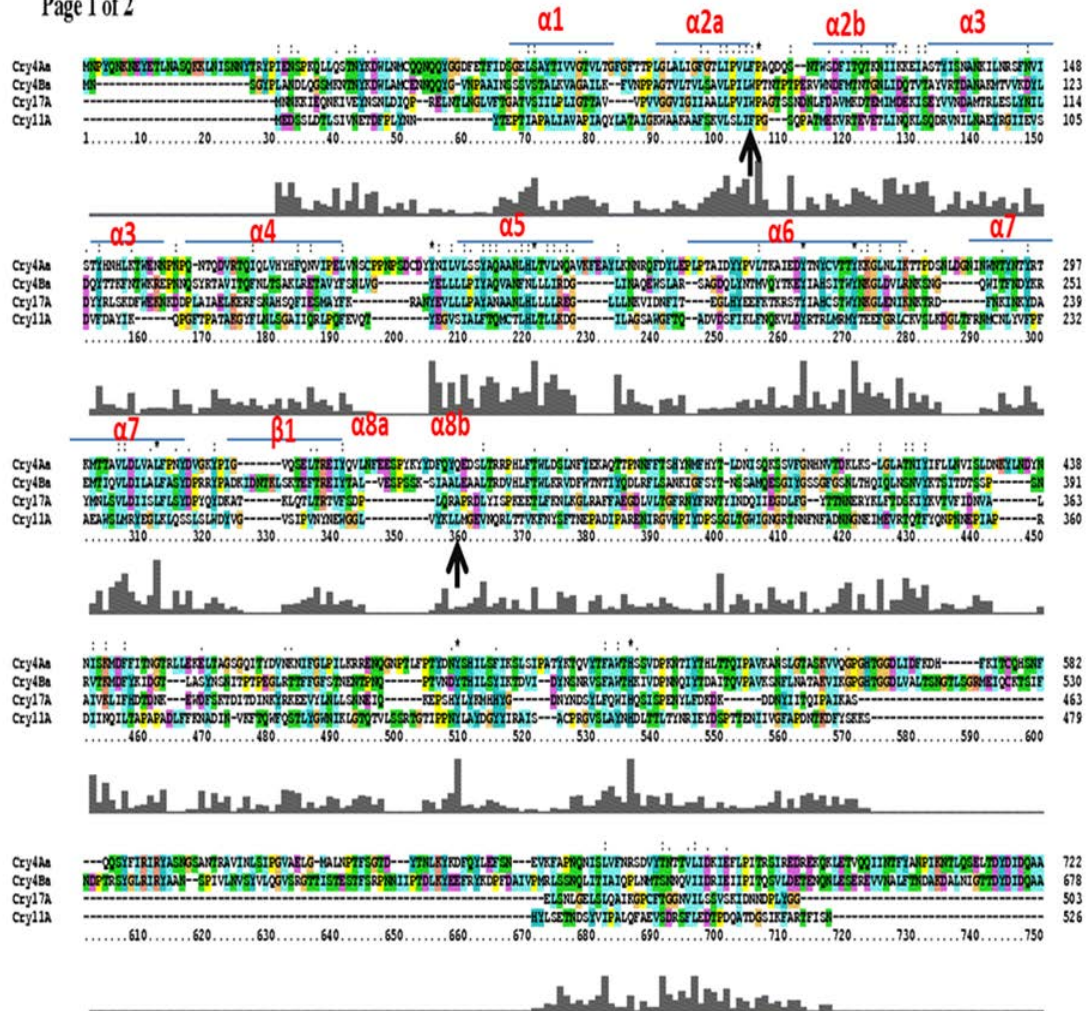region that was deleted by 636bp MscI-ApaI fragment deletion.

Figure 4.6: Multiple sequence alignment of Cry17A with Cry4A, Cry4B and Cry11A. Clustal W analysis of the Cry17A and Cry4Ba protein sequences showed that 636 bp MscI-ApaI fragment deletion of Cry17A gene is a deletion of the whole domain from α2b to α8b, indicated by black arrows. The domains are indicated in red.

A.

Cry17A protein		Cry17A-4B Chimeric protein
71.5	Size in Kda	71.8
618	# of amino acids	624
63	Strongly basic residues	61
77	Strongly acidic residues	69
221	Hydrophobic amino acids	225
194	Polar amino acids	209
5.466	Isoelectric point	5.843
-12.125	Charge at pH=7.0	-6.604
1857	Total bases	1875
26	%GC	28

B.

Figure 4.7: Cry17A-Cry4B chimeric model.

Panel A shows the characteristics of the putative Cry17A-Cry4Ba chimeric protein as compared to the wild type Cry17A.

Panel B shows the homology modelling of the Cry17A-Cry4Ba chimeric protein using Cry4Ba (PDB: 1W99.A) as the template.
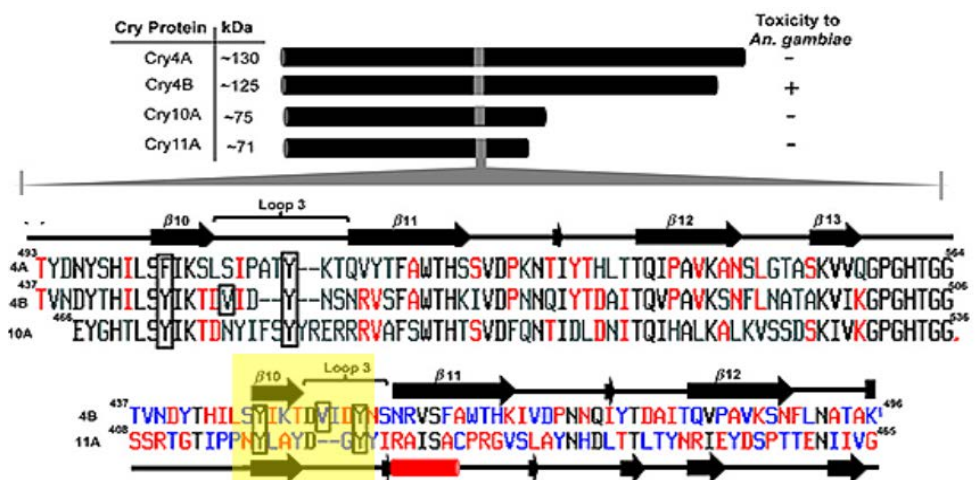
Figure 4.8: Experiment Strategy for Chapter 4. The picture shows predicted sizes of recombinant Cry4A, Cry4B, Cry10A and Cry11A proteins and their toxicity to *An. gambiae* larvae. It also shows the multiple sequence alignment and secondary structure prediction showing the putative receptor-binding regions in Cry4A (residues 493–564), Cry4B (residues 437–506), Cry10A (residues 466–536) and Cry11A (residues 408–465). The predicted secondary structures, β10, loop-3 and β11–β13 strands of Cry4A, Cry4B and Cry10A are indicated above the sequences. The difference between toxic Cry4B and non-toxic Cry11A against *Anopheles gambiae* is the presence to Y----V—Y in Cry4B and Y-----Y in Cry11A (as shown in β10, loop-3 shaded region). (Taken from Ibrahim et al., 2013). The experiment strategy was:

- Introduction of *Anopheles* toxicity into Cry Operon
  - In Cry17A   : 414- YKMHHYGDNY – 424
  - In Cry4B   : 447- YIKTDVIDYN -458
  - Mutate Tyr-420 to Valine (UAU →GUU)

- Introduction of *Culex* toxicity into Cry Operon
  - In Cry4Ba, Loop 3 was mutated to mimic loop3 of Cry4Aa which is toxic to *Culex*. When D454 was replaced with P and AT was inserted after it, caused the toxin to gain activity more than 219-fold against Culex, relative to native Cry4Ba.
  - In Cry17A, I will need to change DNY → PATY.
    Change D →P (AGC → CCC)
    Insert A → GCC
    Change N → T (AAC → ACC)

174

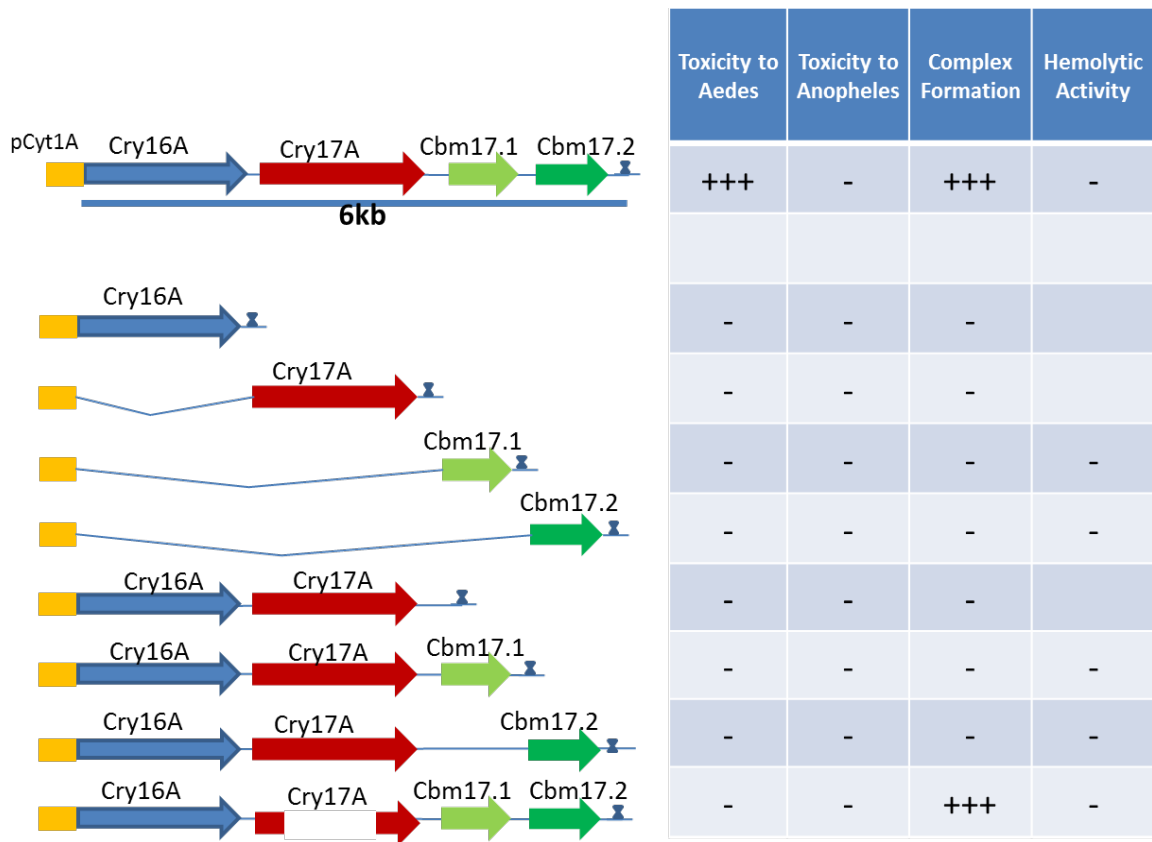|	Toxicity to Aedes	Toxicity to Anopheles	Complex Formation	Hemolytic Activity
pCyt1A–Cry16A–Cry17A–Cbm17.1–Cbm17.2 (6kb)	+++	-	+++	-
Cry16A	-	-	-	
Cry17A	-	-	-	
Cbm17.1	-	-	-	-
Cbm17.2	-	-	-	-
Cry16A Cry17A	-	-	-	
Cry16A Cry17A Cbm17.1	-	-	-	-
Cry16A Cry17A Cbm17.2	-	-	-	-
Cry16A Cry17A Cbm17.1 Cbm17.2	-	-	+++	-

Figure 4.9: Any gene deletion in the Cry operon results in loss of toxicity. Four gene deletions were made. In the first, pCry16, only Cry16A was expressed, while in the second, CryO/Δ17.1, only Cbm17.1 was deleted, in the third, pCry16/17, both Cbm17 genes were deleted and in the fourth, CryO/Δ17, only Cry17 was deleted. The latter three were derived from the pCryO by restriction with PacI (P), NcoI (N),and MscI-ApaI respectively. Each of the constructs is under the control of the Cyt1A promoter. Both CryO and CryO/Δ17 show complex formation but only CryO shows toxicity to *Aedes*.
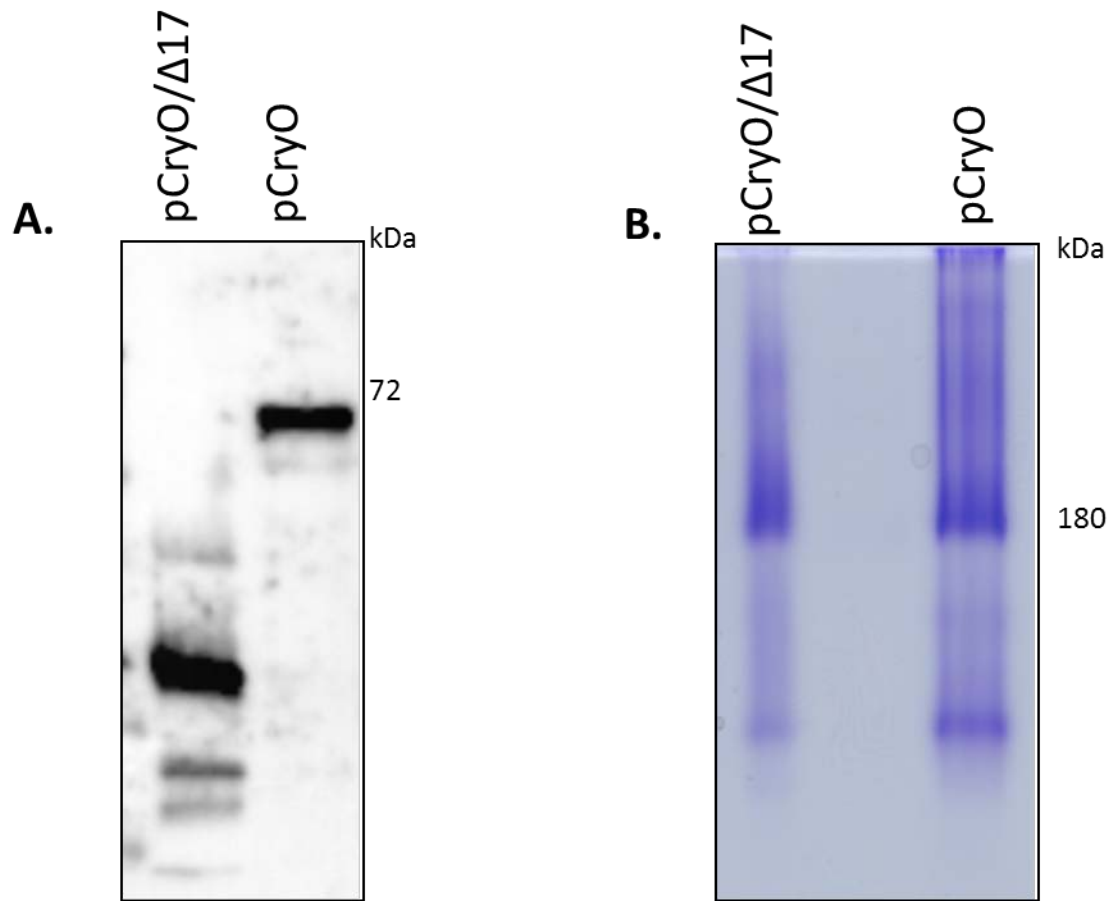
Figure 4.10. Native PAGE shows the CryO/Δ17 forms a complex.

A. The western blot of SDS-PAGE gel shows that entire Cry operon (pCryO) is stably expressed at 18h whereas the Cry17A deletion mutant (pCryO/Δ17) shows that is unstable and degraded.

B. The Native PAGE shows that the entire Cry operon (pCryO) forms a complex of about 170kDa in 24 h rBTcultures. This complex is also shown in the Cry17A deletion mutant (pCryO/Δ17).

Table 4.1: List of sequences of primers used in site-directed mutagenesis.

#	Primer Name	Primer Sequence
1	cry17_Anoph_FW	TCATTATTTATATAAAATGCATCACGTTGGTGATTACAATAATGATTCTTATTTATTCCAATGGAT
2	cry17_Anoph_RV	ATCCATTGGAATAAATAAGAATCATTATTGTAATCACCAACGTGATGCATTTTATATAAATAATGA
3	cry17_Culex_FW	ATGCATCACTATGGTCCTGCCACCTATAATGATTCTTA
4	cry17_Culex_RV	TAAGAATCATTATAGGTGGCAGGACCATAGTGATGCAT

Authors Note:

Stefani Andrews create the pCryO/Δ17 using the MscI-ApaI restriction digest to delete

the 636bp MscI-ApaI fragment from within the Cry17 gene.

Chapter 5

Investigate the mosquitocidal activity of Clostridial Mosquitocidal Protein Toxin Operon of *Clostridium bifermentans* subsp. *malaysia* to *Aedes* and *Anopheles* larvae

Abstract

It has been shown previously that four toxins encoded by the Cry operon, Cry16A, Cry17A, Cbm17.1, and Cbm17.2, are all required for toxicity, and these toxins collectively show remarkable selectivity for *Aedes* rather than *Anopheles* mosquitoes, even though *C. bifermentans* subsp. *malaysia* is more toxic to *Anopheles*. Here we show the toxins that target *Anopheles* are encoded in a new Clostridial Mosquitocidal Protein (Cmp) operon. The Cmp operon encodes eight genes, namely P-47, NTNH, OrfX1, OrfX2, OrfX3, HA and a 150-kDa mosquitocidal toxin (CMP1) All the genes show complete or partial homology to the *Clostridium botulinum* type B neurotoxin cluster genes. We show that the three genes, *CMP1, CNTNH,* and *CHA* that play an important role in botulinum neurotoxicity are not active against mosquitoes.

Introduction

With the aim to identify new strains of insecticidal bacteria, a nationwide screening program in Malaysia led to the description, of the first anaerobic isolate, a *Clostridium bifermentans* strain having a high mosquitocidal activity (Lee and Seleena, 1990a). This strain was isolated from a mangrove swamp soil and identified as *C. bifermentans* serovar *malaysia* (de Barjac et al., 1990; Lee and Seleena, 1990a). *C. bifermantans s*ubsp. *malaysia* (Cbm) is highly toxic to *Anopheles,* with lower toxicity to *Aedes aegypti* and *Culex* (Lee and Seleena, 1990b; Thiery et al., 1992). The biocontrol of mosquitoes for the past decade has been dominated by *Bacillus thuringiensis* subsp. *israelensis* (Bti) toxins. Cbm toxicity against *A. stephensi* is about 10 times higher, but its toxicity against *Culex pipiens* or *Aedes aegypti* is 10 times lower, than that of Bti (Thiery et al., 1992b). Cbm cells show no toxicity to mammals and goldfish (Thiery et al., 1992). Other studies for Cbm toxicity against mammals and various other non-target organisms have also shown a lack of toxicity (Yiallouros et al., 1994). Genetic analysis of putative insecticidal factors of *C. bifermentans* subsp *malaysia* identified four genes of interest. Two genes were structurally related to δ-endotoxin genes of *Bacillus thuringiensis* and named as Cry16A and Cry17A. The other two genes are variants of the same gene and showed 45% similarity with *Aspergillus fumigatus* were named as Cbm17.1 and Cbm17.2. The toxic activity of Cry16A was investigated by expressing Cry16A in a non-toxic Bt strain (Barloy et al., 1996, 1998). Low level of toxicity was seen against second-instar *An. stephensi* larvae, which might have been due to the transformants producing very small amounts of Cry16A protein. With similar protocols, Cry17A showed

negligible activity and Cbm17.1 and Cbm17.2 showed no larvicidal activity and hemolytic activity (Barloy et al., 1998). Later Juárez – Pérez and Delécluse found no mosquitocidal activity in recombinant Cry16A, Cry17A, Cbm17.1 and Cbm17.2, and concluded that mosquitocidal activity of Cbm is due to otherwise unknown factors (Juárez – Pérez et al., 2001).

Bt synthesizes a vast number of protein toxins with activity against a wide range of organisms in nature, including not only a broad range of insect orders but also nematodes, a human-pathogenic protozoan, animal and human parasites plus different human-cancer cell lines (Schnepf et al., 1998; Van Frankenhuyzen, 2009, 2013; Ohba et al., 2009; Kondo et al., 2002; Peña, et al., 2013). One of the reasons driving such wide toxin diversification is the organization of toxin-coding genes in the genome. Most Bt strains bear large plasmids containing their toxin-coding gene repertoires (De Maagd et al., 2003; Agaisse, et al., 1995; Loeza-Lara et al., 2005; Mesrati et al., 2005). Multiple toxin genes are present in the same or different plasmids in a single Bt cell. For example, Bti contains Cry4A, Cry4B, Cry10A, Cry11A and Cyt1Aa toxin genes, all of which are present on one 110-kb plasmid (Ward et al., 1984; Thorne et al., 1986; Berry et al., 2002; Angsuthanasombat et al., 1987; Yamamoto et al., 1988). The movement of plasmids in nature may account for the discovery of cry toxin genes, related to those of Bt, in Ls (Jones et al., 2007), *Paenibacillus popilliae* (Zhang et al., 1997), *Paenibacillus lentimorbus* (Yokoyama et al., 2004) and *Clostridium bifermentans* serovar *malaysia* (Barloy et al., 1996) as well as the presence of cyt-like genes in *Erwinia* (Rigden, 2009).

Clostridium itself is a large and diverse genus, characterized by gram-positive bacteria capable of forming endospores. Most species are well known for their metabolic properties and distinctive industrial applications. However, a small portion is responsible for causing human and animal diseases. Only a small portion is responsible for causing human and animal diseases. This discordant phylogeny between the different metabolic activities, toxin types and the host bacteria indicates that horizontal gene transfer of the genes is responsible for the toxin/metabolite presence within the different bacteria (Rummel and Binz, 2013). It is highly probable that Cbm acquired the Cry Operon (Qureshi et al., 2014) and the Clostridial Mosquitocidal Operon (Cmp Operon) on the pClosMP through horizontal gene transfer.

Our lab has demonstrated that the Cry operon encodes four genes, all of which are required as a complex for toxicity to mosquito larvae (Qureshi et al., 2014). The toxin complex, however, showed surprising species selectivity, particularly to *Aedes* mosquitoes. In this work, we report the cloning and characterization of another operon, the clostridial mosquitocidal protein operon (Cmp), and investigate its mosquitocidal activity against *Aedes*, *Anopheles* and *Culex* mosquitoes. The Cmp operon encodes for eight genes. These genes show homology to the genes that encode for *Clostridium botulinum* type B neurotoxin operon.

Material and Methods

Bacterial strains and culture conditions

C. bifermentans (ATCC) was used as the wild type reference strain and *C. bifermentans* serovar *malaysia* (*Cb malaysia*) was from the collection of the Institute for Medical Research, Malaysia (Lee and Seleena, 1990b). These strains were used for bioassays and for isolation of gDNA for cloning. *Escherichia coli* strainS DH10 beta electrocompetent cells (New England Biolabs, Ipswich, MA) and BL23(DE3) chemical competent cells (Novagen,) were used for cloning and Top10 Cells (Invitrogen, Grand Island, NY) were used for subcloning in TOPO pCR-2.1 vector (Invitrogen, NY). NEB transformation protocol was used. *B. thuringiensis* subsp. *israelensis* 4Q7 (Bacillus Stock Center, Ohio State University, Columbus, Ohio) was used for expression of toxin genes using a published protocol (Macaluso and Mettus, 1991). Clostridial strains were streaked for isolation on TYG agar and liquid cultures were grown in TYG medium at 30°C under anaerobic conditions using BD GasPakEZ (Beckton-Dickinson Microbiology, Cockeysville, MD). The liquid cultures for *Clostridium* strains were grown anaerobically at 30°C, while the *Bacillus* cultures were grown at 30°C with shaking. Nutrient broth media was used for *B. thuringiensis* and Luria broth media was used for *E. coli*. Ampicillin (100µg/ml), kanamycin (50µg/ml) and erythromycin (25µg/ml) were added when required. The bacterial strains were individually grown at 37°C and the His-tagged and S-tagged recombinant proteins were induced by adding 1 mM isopropyl β-D-thiogalactoside (IPTG) for 5 h at 37°C.

General DNA Techniques

All DNA manipulations were performed according to standard protocols. Primers used in this study were created using the sequences from existing databases. Polymerase chain reactions (PCR) were performed in an automated thermacycler (C 1000 Touch ™, Bio Rad). Choice-TaqTM mastermix DNA polymerase (Denville Scientific, Metuchen, NJ) was used for all PCR reactions for products below 2kb. Advantage 2x polymerase (Clontech, Mountain View, CA) or Phusion polymerase (NEB) was used for all PCR reactions for products above 2kb. PCR products were separated in 1% agarose gels and subsequently cut and purified using Wizard SV Gel and PCR purification kits (Promega, Madison, WI). All the Restriction digests were carried in a solution (20 µl) containing 500ng DNA sample, 1 x NEB buffer, and 1 µl each restriction enzyme(s) using NEB recommended protocols. After the reaction mixture was separated in a DNA gel, the target band was cut and purified from the gel with Wizard SV Gel and PCR purification kits (Promega, Madison, WI). These reactions were ligated in 20 µl ligation solution containing 5 µl samples, 1X T4 DNA Ligase Reaction Buffer, and 1µl T4 DNA ligase (New England Biolabs, Ipswich, MA) overnight at 14°C. All PCR products were cloned into the pCR2.1 vector (Invitrogen, Grand Island, NY). Sequencing of purified DNA products was performed by the Genomics Core facility at the University of California, Riverside.

Construction of plasmids expressing the individual toxins CMP1, CNTNH and CHA in *Bacillus* expression system

The vector pHT315 (Arantes and Lereclus, 1991) was used for construction of expression plasmids (Table 5.1). The Cyt1A promoter a sporulation promoter from Bti was used for expression of all pHT315 constructs. Plasmids, pbCmp (*CMP1*), pbCNTNH (*CNTNH*) and pbCHA (*CHA*), the genes were synthesized (Genscript, Piscataway, NJ) and cloned in pHT315 and pUC57. Plasmids were then transformed in DH10 beta electro-competent cells (NEB). Following confirmation of correct sequence, the plasmid constructs (Table 1) were extracted from DH10 beta cells. All pHT315 plasmids were independently transformed *Bt israelensis* 4Q7 cells and by electroporation and colonies isolated for erythromycin resistance as previously described (Chang et al., 1993).

Construction of plasmids expressing the individual toxins CMP1, CNTNH and CHA in *E. coli* expression system

The ampicillin resistant pETDuet™-1and kanamycin resistant pRSFDuet™-1 vectors (Novagen) were used for construction of expression plasmids (Table 5.1). The IPTG inducible T7 promoter was used for expression of all gene constructs in BL21DE3 cells (Novagen). The genes of interest were amplified using primers 1 and 2 for pET-CHA-Stag (CHA) and primers 3 and 4 for pRSF-CMP1 (CMP1)(Table 5.2). The pETDuet™-1was linearizes with NdeI and XhoI restriction enzymes. pRSFDuet™-1 vector was linearized using NcoI in the multiple cloning site. All DNA products were gel purified and purified using Wizard SV Gel and PCR (Promega) purification kits to

provide the desired DNA concentrations. Sub-cloning was done using pCR2.1 vector.

After confirmation of the correct sequence, the positive clones were digested and

fragments ligated. The correctly ligated clones were transformed in DH10 beta cells

(NEB). The plasmid constructs (Table 5.1) were extracted from DH10 beta cells in ample

quantities. Plasmids, pET-CHA-Stag and pRSF-CMP1, were made using this approach.

However, for plasmids, peCMP1 (CMP1-His), peCMP1-CNTNH (CMP1-His, CNTNH-

Stag) and peCHA (CHA-His), the genes were synthesized (Genscript, Piscataway, NJ)

and cloned in pETDuet™-1and pRSFDuet™-1 vectors for us by Genscript (Table 5.1).

All plasmids were then used to independently transform BL21DE3 cells (NEB) by

electroporation and colonies isolated for Kanamycin and Ampicillin resistance as

described in manufacturer protocol.

Homology modeling of Cmp operon toxins

SWISS-MODEL (Arnold et al., 2006) was primarily used for modelling purposes.

It is freely available at http://www.expasy.org/swissmod/SWISS-MODEL.html. Multiple

sequence alignments were performed using CLUSTALW. Models were constructed using

the "Automated (project) mode" in SWISS-MODEL, and viewed with Swiss-Pdb

Viewer. The template file used was chosen from either the known tertiary structures or

from the models that were successfully constructed with SWISS-MODEL. Models of the

eight Cmp operon proteins were constructed with this technique. This protocol did not

always resulted in a model.

Bioassays

Total cultures of *C. bifermantans*, *Cb* malaysia, recombinant *E. coli* (rEC)and recombinant *B. thuringiensis* (rBT) strains were assayed against *Ae. aegypti* and *An. gambiae* larvae. Recombinant *E. coli* bacterial strains were individually grown at 30°C till O.D.$_{600}$ of 0.6 and the N-terminal His-tagged and C-terminal S-tagged recombinant proteins were induced by adding 1 mM isopropyl β-D-thiogalactoside (IPTG) for 18 h at 30°C. Bioassays were performed at 24° C using 25 third instar larvae in 200 ml water. Live bacterial cultures were used in bioassays. Larval mortality was determined by counting the number of larvae still alive at 24 and 48h. Bioassays were repeated 2-3 times and LC50 concentrations were determined by probit analysis (USDA) and plotted using the Origin program (Origin Lab, Northampton, MA).

Polyacylamide gel electrophoresis (PAGE) and immunoblot

C. bifermantans, *Cb malaysia*, rBT , rEC, BL21DE3 and 4Q7 strains were grown in parallel for 18h. Aliquots, 100ul, of the cultures were mixed with of 2X sample buffer (Bio-Rad, Hercules, CA), and the samples separated by SDS-polyacrylamide (10-14%) gel electrophoresis. Native polyacrylamide (10%) gel electrophoresis was performed using 1X native PAGE buffer (3g Tris, 14.5g glycine). In parallel the separated proteins were transferred onto PVDF membranes (Millipore, Billerica, MA). The membranes were treated with blocking solution (1X PBS, 5% skim milk and 0.1% Tween-20) for 1 h at room temperature and then washed with PBST (1XPBS and 0.1% Tween-20). The blocked membrane was incubated overnight (at 4°C) with primary anti-CMP1 and anti-

His-tag at 1:3000 dilution . A secondary antibody, ECL™ anti-rabbit IgG-horseradish peroxidase-linked whole antibody (GE Healthcare, Anaheim, CA) was used at a dilution of 1:5000. Immunoreactive bands were visualized using the ECL western blotting kit (GE Healthcare) and exposed to an X-ray film.

Results

Structure of Clostridial Mosquitocidal Protein Operon.

Botulinum neurotoxin is produced as a non-covalently bound complex of two or more protein components. The complex can include hemagglutinin proteins (HA) as well as a non-toxic non hemagglutinin protein (NTNH). The genes encoding these non-toxin proteins are clustered with that encoding BoNT. The *NTNH* gene is located immediately upstream of that encoding BoNT in all toxin types, and genes for other components of complexes have been shown to be clustered upstream of the *NTNH* gene. Almost always the cluster is associated with either HA genes (*HA17, HA33, HA70*) or *OrfX* genes (*OrfX1, OrfX2, OrfX3*) or *P-47* gene but never all together or in combination (Collins and East 1998; Rummel and Binz, 2013).

In Cbm, we have a unique organization of the Cmp operon on the pClosMP plasmid. It is present downstream of the Cry operon, flanked by insertion sequences and transposon elements on both ends in what looks like a type IV pilus (Figure 5.1). The Cmp operon contains all the non-toxic components of the BoNT cluster except P-21 gene. A single *HA* gene is a separate transcription unit (Figure 5.2). In the opposite

orientation *CMP1, OrfX1, OrfX2, OrfX3* and *NTNH* genes are present as an operon in the stated order, followed by *P-47* genes as a separate transcription unit. The Cmp operon genes show homology to, Botulinum neurotoxin B/E gene, *OrfX1* gene, *OrfX2* gene, *OrfX3* gene, *NTNH* gene and transcripHA and *P-47* gene (Table 5.3). Since genes encoding these complex components are physically adjacent to genes responsible for BoNT production (as in *C. botulinum*), it is not unreasonable to expect that their genealogical relationships would mirror that of the neurotoxin. The orientation of these genes is different but the RBS and promoter regions for each gene and operon suggest that these proteins should be expressed in Cbm.

Expression of the Clostridial Mosquitocidal Protein operon

The proteins reacting with CMP1 antibodies were observed in whole cultures of Cbm, but not in Cb cultures. It was seen that CMP1 is expressed as a 140kDa protein (Figure. 5.8A), and that it forms a complex of approximately 480kDa (Figure 5.8B) with other proteins in the Cbm culture. A nonspecific 55kDa clostridial protein reacting with CMP1 antibodies was observed in whole cultures of Cbm, Cbp Cb but not the 4Q7 containing the pHT315 vector alone (Fig. 5.8A).

Individual genes of Cmp operon, namely, *CMP1, CNTNH* and *CHA* genes were expressed using the Cyt1A promoter from *Bt israelensis* (Chang et al., 1993; Wu and Federici, 1993) in pHT315 and T7 promoter in Duet vectors (Novagen). Protein expression from all *Bacillus* and *E. coli* transformations, rBT and rEC respectively, were analyzed by Coomassie Blue staining and western blotting. For rEC strains (Figure 5.7),

the cultures were grown after adding 1mM IPTG for 4, 12, 18, 24 and 48h at 30°C to determine the optimal time for toxin production in *E.coli*. In western blot, a 140 kDa corresponding to the predicted size of the CMP1 toxin, and one at about 51kDa, which corresponds to the CHA was detected by the anti-6x-His Tag antibody but the S-tag antibody could not detect any band for the CNTNH protein. No protein reacting with CMP1, His-tag and S-tag antibodies were observed in control cultures of 4Q7 containing the pHT315 vector alone and BL21DE3 cells containing pETDuet™-1and pRSFDuet™-1 alone. Thus we concluded the three genes are expressed under different conditions in rBT and rEC strains, with an optimal expression of rEC at 30°C.

rBT cultures were grown for 18h at for all pHT315 strains. Strong bands were observed around 140kDa for CMP1 but no expression was observed in CNTNH and CNTNH-HA 24 h cultures (Figure 5.6A). This could be due to the possibility that the antibody synthesized against CNTNH did not work properly. These bands were not present in the control comprising of 4Q7 strain containing the pHT315 vector alone.

Larvicidal activity of the individual Clostridial Mosquitocidal Protein Operon genes expressed in *Bacillus* expression system.

Whole cell preparations from *C. bifermantans*, *Cb malaysia*, and rBT transformants containing pHT315 (control) and pbCMP1, pbCNTNH and pbCNTNH-CHA (Table 5.1) that were expressed were assayed for larvicidal activity using third instar *Ae. aegypti* and *An. gambiae* (Figures 5.9 and 5.10 respectively). The rBT pHT315 control samples showed no larval toxicity against both species. The rBT pbCMP1,

pbCNTNH, pbCNTNH-CHA showed no toxicity against *An. gambiae., Ae. aegypti* and *Culex quinquefasciatus.* Only Cbm cultures showed mortality against *An. gambiae* with an LC_{50} of 0.5µl (Figure 5.10) and *Ae. aegypti* with an LC_{50} of 150µl (Figure 5.9).

Larvicidal activity of the individual Clostridial Mosquitocidal Protein Operon genes expressed in *E. coli* expression system.

      E. coli remains the primary system of choice of obtaining large quantities at low cost and at short time, for either individual proteins or protein complexes. We synthesized peCMP1 with His-tag, peCMP1-CNTNH with S-tag on NTNH and peCHA with His-tag for easier purification of proteins and wanted to test the activity of the proteins. These toxins were cloned under control of the IPTG inducible T7 promoter in Duet vectors with the plan to co-express proteins for better understanding the complex activity. All the plasmids were shown to form a stable protein (Fig. 5.7), but more importantly none of the bacterial strains containing these constructs showed any larval toxicity to either *Anopheles gambiae* or *Aedes aegypti*. Even though the synthesized clones were codon optimized for *E.coli* and expressed, they showed no mosquitocidal activity.

Homology Modelling of Cmp Operon proteins

      We attempted to create models of the seven genes encoded by the Cmp operon and CHA gene. SWISS-MODEL gives all the models a GQME score. GMQE (Global Model Quality Estimation) is a quality estimation which combines properties from the

target-template alignment. The resulting GMQE score is expressed as a number between zero and one, reflecting the expected accuracy of a model built with that alignment and template. Higher numbers indicate higher reliability. SWISS-MODEL gave no model for OrfX1. The models for OrfX2, OrfX3 and P-47 were of very poor accuracy with GMQE value ranging from 0.2 – 0.5 and were discarded. Figure 5.3 depicts the analysis of CHA protein domains with a putative model of CHA based on Glycerophoshphodiesterase (GpdQ) from *Enterobacter aerogenes* (Hadler et al., 2008) and showed coverage in the central and C –terminal part of the protein only. Figure 5.4 shows the analysis of CMP1 protein domains with a putative model of Cmp1 based on Crystal structure of the Botulinum neurotoxin type B complexed with synaptotagamin-II ectodomain (Chai et al., 2006) showed almost 100% coverage. Figure 5.5 depicts the analysis of CNTNH protein domains with a putative model of CNTNH based on crystal structure of BoNT/Ai in complex with NTNHA (Gu et al., 2012).

Discussion

The anaerobic bacterial strain *C. bifermantans malaysia* is highly active against a number of mosquito genera but particularly against anopheline species (de Barjac et al., 1990; Lee and Seleena, 1990b). Prior efforts characterized the bacteria and identified proteins that had mosquitocidal activity (Charles et al., 1990; Nicolas et al., 1990; Nicolas et al., 1993) and thought to have a novel mode of action. The genes encoding these proteins were subsequently identified (Barloy et al., 1996; Barloy et al., 1998b). All

of the identified proteins (Nicolas et al., 1993) are encoded in a single operon, the Cry

operon (Barloy et al., 1996; Qureshi et al., 2014). However, the genes responsible for its

toxicity in *Anopheles* have remained elusive.

Genome sequencing of this bacterium led to the finding of Cmp locus on its

largest plasmid pClosMP along with the Cry operon. This locus consists of seven genes,

CMP1, OrfX1, OrfX2, OrfX3, CNTNH, p47 and *CHA*, which is a separate transcription

unit. The derived sequence of individual Cmp genes suggests that these proteins are like

botulinum toxin cluster family and are similar in sizes (Table 5.3) and are acquired by

Cbm through horizontal gene transfer between other firmicutes such as *Paenibacillus* and

Clostridium species. The major indication of this event is the presence of transposons and

insertion sequences flanking the Cmp operon along with the presence of Type IV pilus

genes. Furthermore, the gene sequences show homology to known toxin genes. The CHA

has about 45% amino acid identity to the partial HA33 proteins, while the OrfX1 has

about 48% identity to the OrfX1 protein, OrfX2 has about 55% identitiy to the OrfX2

protein, OrfX3 has about 67% identity to OrfX3 protein and p47 has about 45% to the

partial P47 protein in *Paenibacillus dendritiformis*. In contrast the CMP1 and CNTNH

proteins identify to the *C. botulinum* type B family of proteins (Berne et. al., 2009). These

two proteins are 32% and 42% similar in amino acid sequence to BoNT/B and NTNH

and are 150kDa and 146kDa in size respectively.

Botulinum toxins are produced in complexes with non-toxic associated proteins

(NAPs). The neurotoxin genes (BoNT) in *C. botulinum* are located in a cluster with the

genes coding for the NAPs in the botulinum locus, and all the genes in the cluster are

expressed together (Couesnon et al., 2006). Each BoNT is encoded by an approximately 3.8 kb gene, which is preceded by a non-toxic non-hemagglutinin gene (NTNH) and several other genes that encode toxin-associated proteins (hemagglutinins HA-17, HA-33, HA-70, OrfX1, OrfX2, OrfX3, p21, and/or p47) (Collins and East, 1998; Kubota et al., 1998; Sharma et al., 2003;). The genes are transcribed as two divergent polycistronic messages, with the NTNH and BoNT in one and HA70, HA17 and HA33 in the other (Marvaud et al., 1998; Oguma et al., 1999)(Figure 5.2). Organizationally, the Cmp operon also contains the same structure but also shows some variation. A gene called botR or cntR, located in the locus, has been found in all types except non-proteolytic type E and acts as a positive regulator (Marvaud et al., 1998) but is not annotated yet in Cmp operon. Cmp operon contains both the p47 and CHA genes in the same cluster and contains only one HA gene homolog.

In addition, NAPs are known to potentiate oral toxicity of the BoNT toxin complex. It was shown that the larger the molecular size of the toxin complex, the higher the oral toxicity (Sharma et al., 2003). NTNH shields BoNT by forming an interlocked complex (Gu et al., 2012) in the harsh conditions in the digestive tract, thereby potentiating the oral toxicity. HA is also reported to contribute to the stability of BoNT by interacting with epithelial cells through carbohydrate binding and disrupts the epithelial barrier by binding to E-Cadherin and disrupting its function (Fujinaga et al. 2013; Lee et al. 2013; Sugawara et al., 2010, 2014).

Since genes encoding Cmp operon components are structurally similar to genes responsible for BoNT production, it is not unreasonable to expect that their genealogical

relationships would mirror that of the BoNT. The ribosomal binding sites (RBS) and promoter regions for each gene in the Cmp locus suggest that these proteins should be expressed in Cbm. Therefore, models of three NAP component genes CMP1, CNTNH and CHA were created using Swiss Model Workspace (Arnold et al., 2006).The models (Figures 5.3, 5.4, 5.5) showed that the N-terminal fragment of the CMP1 and CNTNH proteins contain the functional HELIH and KCLIK sites but the C-terminal fragments that are the binding domains of the toxin proteins are different. This could account for the fact that Cbm toxins are not toxic to vertebrates whereas BoNTs are.

In our efforts to identify novel mosquitocidal genes from Cbm, we investigated the individual genes of these toxins in this Cmp locu, as the cloning of the full length 16kb operon proved to be difficult. We see the expression of the Cmp operon genes in Cbm whole cultures using the Anti-CMP1 antibody (Figure 5.8) at 24 hr but not in the non-mosquitocidal Cb whole culture. As *CMP1, CNTNH,* and *CHA* play an important role in botulinum neurotoxicity, the three genes were individually expressed under control of the Cyt1A promoter. None of these proteins were toxic, although we could confirm only the expression of the full-length protein for only one toxin, CMP1. A distinct and larger complex was also visible in the native PAGE and immunoblot of Cbm whole bacterial culture (Fig. 5.8B).The interaction of BoNT with HA and NTNH in *C. botulinum* suggested that the Cbm homologs CMP1, CHA and CNTNH were expressed simultaneously as a stable complex and act together. Based on this premise we tried to create and express a clone of the CMP1-CNTNH-CHA under the control of a strong

Bacillus promoter, Cyt1A (Chang et al., 1993; Wu and Federici, 1993). However, we were not able to clone it.

Since the purification of tagged proteins is easier in *E. coli* expression system, the three Cmp operon genes were co-expressed under control of the inducible T7 promoter. The expression of CMP1 protein was not very good and could be obtained with difficulty by inducing the culture at 20°C. However, none of these proteins were toxic, although we could confirm the expression of the full-length protein for all three. The reason could be that these clostridial proteins could not be secreted by the *E. coli* expression system. Our results are therefore consistent with previous observations (Nicolas et al., 1993) that Cbm toxicity against *A. stephensi* is higher than its toxicity against *Culex quinquefasciatus* or *Aedes aegypti*.

Cbm is the first known anaerobic bacterium shown to possess larvicidal activity. We have already shown that Cry operon genes, similar to those found in *B. thuringiensis*, previously thought to be toxic to *A. stephensi*, and were actually toxic to *Aedes aegypti*. This raised the question about the factors that were toxic against *A. stephensi*. It is clear the Cmp operon in Cbm contains a very unique combination of toxins that are highly mosquitocidal and nontoxic to humans. It is a likely possibility that Cbm acquired this operon with horizontal gene transfer. The Cry operon shows high selectivity towards *Aedes* mosquitoes. Does Cmp operon show toxicity to *An. gambia*? It is a possibility that individually the genes are not stable and need all the other genes of the operon to be stable and active. Therefore, we need to express the native full length 16kb Cmp operon and CMP1-CNTNH-HA constructs to answer this question.

References

Agaisse, H. and Lereclus, D. (1995). How does *Bacillus thuringiensis* produce so much insecticidal crystal protein? *J. Bacteriol.* **177**: 6027–6032.

Arantes, O. and Lereclus, D. (1991). Construction of cloning vectors for *Bacillus thuringiensis*. *Gene*. **108**: 115-119.

Arnold K., Bordoli L., Kopp J., and Schwede T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*. **22:** 95-201.

Angsuthanasombat, C., Chungjatupornchai, W., Kertbundit, S., Luxananil, P., Settasatian, C., Wilairat, P. and Panyim, S. (1987). Cloning and expression of 130 kDa mosquito-larvicidal delta endotoxin gene of *Bacillus thuringiensis* var. *israelensis* in *Escherichia coli. Mol.Gen.Genet.* **208**:384-389.

Arantes, O. and Lereclus, D. (1991). Construction of cloning vectors for *Bacillus thuringiensis. Gene.* **108**:115-119.

Barloy, F., Delecluse, A., Nicolas, L., and Lecadet, M.M. (1996). Cloning and expression of the first anaerobic toxin gene from *Clostridium bifermentans* subsp. *malaysia*, encoding a new mosquitocidal protein with homologies to *Bacillus thuringiensis* delta-endotoxins. *J Bacteriol*. **78**: 3099-3105.

Barloy, F., Lecadet, M.M., and Delecluse, A. (1998a). Distribution of clostridial cry-like genes among *Bacillus thuringiensis* and *Clostridium* strains. *Curr Microbiol.* **36**: 232-237.

Barloy, F., Lecadet, M.M., and Delecluse, A. (1998b) Cloning and sequencing of three new putative toxin genes from *Clostridium bifermentans* CH18. *Gene*. **211**: 293-299.

Berne S, Lah L, Sepcic K. (2009). Aegerolysins: structure, function, and putative biological role. *Protein Sci*. **18**:694-706.

Berry, C., O'Neil, S., Ben-Dov, E., Jones, A.F., Murphy, L., Quail, M.A., Holden, M.T., Harris, D., Zaritsky, A. and Parkhill, J.; *et al*. Complete sequence and organization of pBtoxis, the toxin-coding plasmid of *Bacillus thuringiensis* subsp. *israelensis*. *Appl. Environ. Microbiol.* **2002**, *68*, 5082–5095.

Chang, C., Yu, Y.M., Dai, S.M., Law, S.K., and Gill, S.S. (1993). High-level cryIVD and cytA gene expression in *Bacillus thuringiensis* does not require the 20-kilodalton protein, and the coexpressed gene products are synergistic in their toxicity to mosquitoes. *Appl Environ Microbiol*. **59**: 815-821.

Chai, Q., Arndt, J.W., Dong, M., Tepp, W.H., Johnson, E.A., Chapman, E.R. and Stevens, R.C. (2006). Structural basis of cell surface receptor recognition by botulinum neurotoxin B. *Nature*. **444**(7122):1096-100.

Charles, J.-F., and Nielsen-LeRoux, C. (2000). Mosquitocidal bacterial toxins: Diversity, mode of action and resistance phenomena. *Mem Institut Oswaldo Cruz*. **95**: 201-206.

Charles, J.F., Nicolas, L., Sebald, M., and de Barjac, H. (1990). *Clostridium bifermentans* serovar *malaysia*: sporulation, biogenesis of inclusion bodies and larvicidal effect on mosquito. *Research Microbiol*. **141**: 721-733.

Collins, M.D. and East, A.K. (1998). Phylogeny and taxonomy of the foodborne pathogen *Clostridium botulinum* and its neurotoxins. *J Appl. Microbiol*. **84**, 5-17.

Crickmore, N., Zeigler, D.R., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J., Bravo, A. and Dean, D.H. (2014). *Bacillus thuringiensis* toxin nomenclature. http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/index.html

de Barjac, H., Sebald, M., Charles, J.F., Cheong, W.H., and Lee, H.L. (1990). *Clostridium bifermentans* serovar *malaysia*, a new anaerobic bacterium pathogen to mosquito and blackfly larvae. *C R Acad Sci III*. **310**: 383-387.

De Maagd, R.A., Bravo, A., Berry, C., Crickmore, N. and Schnepf, H.E. (2003). Structure, diversity, and evolution of protein toxins from spore-forming entomopathogenic bacteria. *Annu. Rev. Genet*. **37**: 409–433.

Fujinaga, Y., Sugawara, Y. and Matsumura, T. (2013). Uptake of botulinum neurotoxin in the intestine. *Curr Top Microbiol Immunol* **364**: 45–59.

Georghiou, G.P., and Wirth, M.C. (1997). Influence of exposure to single versus multiple toxins of *Bacillus thuringiensis* subsp. *israelensis* on development of resistance in the mosquito *Culex quinquefasciatus* (Diptera: Culicidae). *Appl Environ Microbiol*. **63**: 1095-1101.

Gu, S., Rumpel, S., Zhou, J., Strotmeier, J., Bigalke, H., Perry, K., Shoemaker, C.B., Rummel, A. and Jin, R. (2012). Botulinum neurotoxin is shielded by NTNHA in an interlocked complex. *Science* . **335**(6071), 977–981

Hadler, K. S., Tanifum, E. A., Yip, S. H., Mitić, N., Guddat, L. W., Jackson, C. J., Gahan, L. R., Nguyen, K., Carr, P. D., Ollis, D. L., Hengge, A. C., Larrabee, J. A. and Schenk G. (2009). Substrate-promoted formation of a catalytically competent binuclear center and regulation of reactivity in a glycerophosphodiesterase from *Enterobacter aerogenes*. *J Am Chem Soc*.**130**(43):14129-38.

Juarez-Perez, V., and Delecluse, A. (2001). The Cry toxins and the putative hemolysins of *Clostridium bifermentans* serovar *malaysia* are not involved in mosquitocidal activity. *J Invertebr Pathol*. **78**: 57-58.

Jones, G.W., Nielsen-Leroux, C., Yang, Y., Yuan, Z., Dumas, V.F., Monnerat, R.G. and Berry, C. (2007). A new Cry toxin with a unique two-component dependency from *Bacillus sphaericus*. *FASEB J*. **21**: 4112–4120.

Kondo, S., Mizuki, E., Akao, T. and Ohba, M. (2002). Antitrichomonal strains of *Bacillus thuringiensis*. *Parasitol. Res*. **88**:1090–1092.

Lacey, L. A. (2007). *Bacillus thuringiensis* serovariety *israelensis* and *Bacillus sphaericus* for mosquito control. *J Am Mosq Control Assoc* **23**: 133-163.

Lederberg, J., Shope R.E., Oakes S.C. (1992). Emerging infections: microbial threats to health in the United States. *Washington: National Academy Press*

Lee, H.L., and Seleena, P. (1990a). Isolation of indigenous larvicidal microbial control agents of mosquitos: the Malaysian experience. *Southeast Asian J Trop Med Public Health*. **21**: 281- 287.

Lee, H.L., and Seleena, P. (1990b). Isolation and evaluation of larvicidal *Clostridium bifermantans* against mosquitoes of public health importance. *Trop Biomed*. **7**: 103-106.

Lee, K., Gu, S., Jin, L., Le, T.T. and Cheng, L.W. (2013). Structure of a bimodular botulinum neurotoxin complex provides insights into its oral toxicity. *PLoS Pathog* **9**: e1003690.

Loeza-Lara, P.D., Benintende, G., Cozzi, J., Ochoa-Zarzosa, A., Baizabal-Aguirre, V.M., Valdez-Alarcón, J.J. and López-Meza, J.E. (2005). The plasmid pBMBt1 from *Bacillus thuringiensis* subsp. *darmstadiensis* (INTA Mo14–4) replicates by the rolling-circle mechanism and encodes a novel insecticidal crystal protein-like gene. *Plasmid* **54**: 229–240.

Macaluso, A., and Mettus, A.M. (1991). Efficient transformation of Bacillus thuringiensis requires nonmethylated plasmid DNA. *J Bacteriol*. **173**: 1353-1356.

Maharaj, R., Morris, N., Seocharan, I., Kruger, P., Moonasar, D., Mabuza, A., Raswiswi, E. and Raman, E. (2012). The feasibility of malaria elimination in South Africa. *Malaria Journal*. **11**:423.

Murray, C.J., Rosenfeld, L.C., Lim, S.S., Andrews, K.G., Foreman, K.J., Haring, D. (2012). Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet*. **379**: 413- 431.

Murray NE, Quam MB, Wilder-Smith A. (2013). Epidemiology of dengue: past, present and future prospects. *Clin Epidemiol* **5:**299-309.

Mesrati, L.A., Tounsi, S. and Jaoua, S. (2005). Characterization of a novel *vip3*-type gene from *Bacillus thuringiensis* and evidence of its presence on a large plasmid. *FEMS Microbiol. Lett.* **244**: 353–358.

Najera, J.A. (2001). Malaria control: achievements, problems and strategies. *Parassitologia*. **43**: 346 1-89.

Nicolas, L., Hamon, S., Frachon, E., Sebald, M., and Barjac, H. (1990). Partial inactivation of the mosquitocidal activity of *Clostridium bifermentans* serovar *malaysia* by extracellular proteinases. *Applied Microbiol. Biotechnol.* **34**: 36-41.

Nicolas, L., Charles, J.F., and de Barjac, H. (1993). *Clostridium bifermentans* serovar *malaysia*: characterization of putative mosquito larvicidal proteins. *FEMS Microbiol Lett.* **113**: 23-28.

Nielsen-Leroux, C., Charles, J.F., Thiery, I., and Georghiou, G.P. (1995). Resistance in a laboratory population of *Culex quinquefasciatus* (Diptera: Culicidae) to *Bacillus sphaericus* binary toxin is due to a change in the receptor on midgut brush-border membranes. *Eur J Biochem*. **228**: 206-210.

Ohba, M., Mizuki, E. and Uemori, A.(2009). Parasporin, a new anticancer protein group from *Bacillus thuringiensis*. *Anticancer Res.* **29**: 427–433.

Peña, G., Aguilar Jiménez, F.A., Hallal-Calleros, C., Morales-Montor, J., Hernández-Velázquez, V.M. and Flores-Pérez, F.I. (2013). *In vitro* ovicidal and cestocidal efects of toxins from *Bacillus thuringiensis* on the canine and human parasite *Dipylidium caninum*. *BioMed Res. Int.* 174619. doi:10.1155/2013/174619.

Qureshi, N., Chawla, S., Likitvivatanavong, S., Lee, H. L., and Gill, S.S. (2014). The cry toxin operon of *Clostridium bifermentans* subsp. *malaysia* is highly toxic to *Aedes* Larval Mosquitoes. *Appl Environ Microbiol*. **80**(18):5689-97.

Rao, D.R., Mani, T.R., Rajendran, R., Joseph, A.S., Gajanana, A., and Reuben, R. (1995). Development of a high level of resistance to *Bacillus sphaericus* in a field population of *Culex quinquefasciatus* from Kochi, India. *J Am Mosq Control Assoc*. **11**: 1-5.

Rigden, D.J. (2009). Does distant homology with Evf reveal a lipid binding site in *Bacillus thuringiensis* cytolytic toxins? *FEBS Lett.* **583**: 1555–1560.

Roberts, D.R., Manguin, S., and Mouchet, J. (2000). DDT house spraying and re-emerging malaria. *Lancet.* **356**: 330-332.

Rummel, A and Binz, T. (2013). Botulinum Neurotoxins. *Current Topics in Microbiology and Immunology* **364**, DOI: 10.1007/978-3-642-33570-9_1, *Springer-Verlag Berlin Heidelberg*.

Schnepf, E., N. Crickmore, J. VanRie, D. Lereclus, J. Baum, J. Feitelson, D.R. Zeigler, and D. H. Dean. (1998). *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev*. **62**:775-806.

Staples, J.E., Gershman, M., and Fischer, M. (2010). Yellow fever vaccine: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recommendations and reports : Morbidity and mortality weekly report Recommendations and reports. *Centers for Disease Contro.* **59**: 1-27.

Sugawara, Y., Matsumura, T., Takegahara, Y., Jin, Y. and Tsukasaki, Y.(2010). Botulinum hemagglutinin disrupts the intercellular epithelial barrier by directly binding E-cadherin. *J Cell Biol* **189**: 691–700.

Sugawara, Y., Yutani, M., Amatsu, S., Matsumura, T. and Fujinaga, Y. (2014). Functional Dissection of the *Clostridium botulinum* Type B Hemagglutinin Complex: Identification of the Carbohydrate and E-Cadherin Binding Sites. *PLoS ONE* **9(10)**: e111170.

Tabashnik, B. E., Gassmann, A. J., Crowder, D. W. and Carriere, Y. (2008). Insect resistance to Bt crops: evidence versus theory. *Nat Biotechnol* **26**: 199-202.

Thiery, I., Hamon, S., Gaven, B., and De Barjac, H. (1992) Host range of *Clostridium bifermentans* serovar *malaysia*, a mosquitocidal anaerobic bacterium. *J Am Mosq Control Assoc*. **8**: 272-277.

Thorne, L., Garduno, F., Thompson, T., Decker, D., Zounes, M.,Wild, M., Walfield, A.M. and Pollock, T. (1986). Structural similarity between the Lepidoptera- and Diptera-specific endotoxin genes of *Bacillus thuringiensis* subsp. *kurstaki* and *israelensis. J. Bacteriol.* **166**: 801-811.

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

Ward, E.S., Ellar, D.J. and Todd, J.A. (1984). Cloning and expression in *Escherichia coli* of the insecticidal delta-endotoxin gene of *Bacillus thuringiensis* var. *israelensis. FEBS Lett.* **175**: 377- 385.

World Health Organization. (2013). World Health Report. WHO, Geneva.

WHO Malaria Report. (2013). World Health Report. WHO, Geneva.

Wu, D., and Federici, B.A. (1993) A 20-kilodalton protein preserves cell viability and promote CytA crystal formation during sporulation in *Bacillus thuringiensis*. *J Bacteriol.* **175**: 5276- 5280.

Yamamoto, T., Watkinson, I. A., Kim, L.,Sage, M. V., Stratton, R., Akanda, N.,Li, Y., Ma, D. P. and Roe, B. A. (1988). Nucleotide sequence of the gene coding for a 130-kDa mosquitocidal protein of *Bacillus thuringiensis israelensis*. *Gene*. **66**: 107-120.

Yiallouros, M., Storch, V., Thiery, I., and Becker, N. (1994). Efficacy of *Clostridium bifermentans* serovar *malaysia* on target and nontarget organisms. *J Am Mosq Control Assoc*. **10**: 51-55.

Yokoyama, T., Tanaka, M. and Hasegawa, M. (2004). Novel cry gene from *Paenibacillus lentimorbus* strain *semadara* inhibits ingestion and promotes insecticidal activity in *Anomala cuprea* larvae. *J. Invertebr. Pathol.* **85**: 25–32.

Zhang, J., Hodgman, T.C., Krieger, L., Schnetter, W. and Schairer, H.U. (1997). Cloning and analysis of the first cry gene from *Bacillus popilliae*. *J. Bacteriol.* **179**: 4336–4341.

Zaim, M. and Guillet, P. (2002). Alternative insecticides: an urgent need. *Trends Parasitol*. **18**(4):161-3.
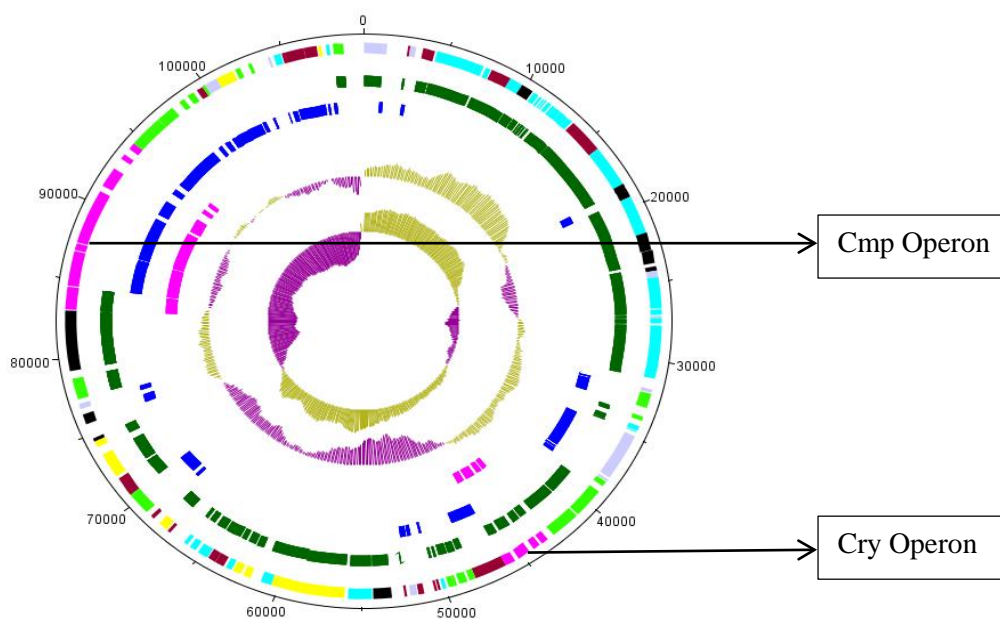
Figure 5.1: Circular representation of 109 Kb plasmid pClosMP. The inner circle represents GC bias [(G − C)/(G + C)], with positive values in Beige and negative values in purple; The second circle represents G+C content; The third circle represents the toxin containing operons (pink); The fourth circle represent predicted genes on the forward strand(dark blue); The fifth circle represent predicted genes on the reverse strand (dark green) and the sixth strand represents all the predicted genes with the following color coding:

gray =   regulatory;
pink =   toxin operons;
blue =   conserved hypothetical;
red =   unknown;
green = transposon related ; mobile elements;
Black = cell wall associated; surface associated;
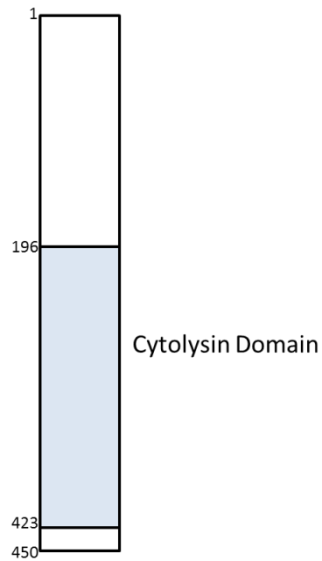yellow = miscellaneous metabolic genes.
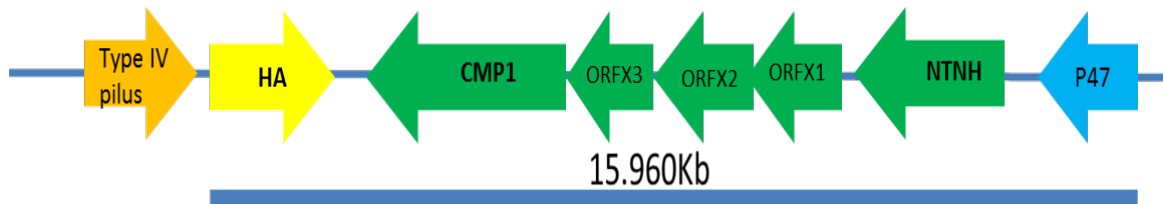
Figure 5.2: Schematic representation of Cmp Locus showing organization of genes.
Schematic representation of the topology and orientation of the Cmp operon(green) and
HA(yellow) and P-47(blue) gene transcription unit as present on the pClosMP plasmid in
Cbm.

Glycerophoshphodiesterase (GpdQ)
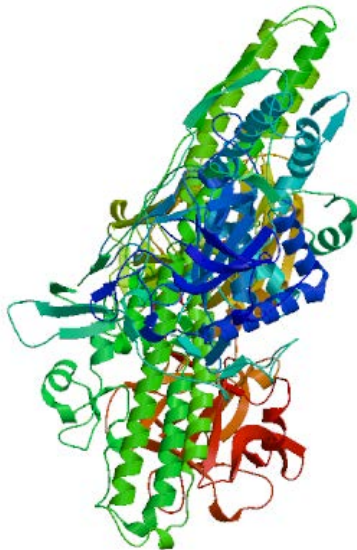from *Enterobacter aerogenes*
PDB: 3d03.1.A
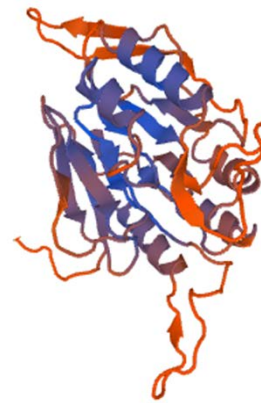
CHA model based on PDB: 3d03.1.A
Sequence Similarity: 16.88%
GMQE score: 0.29

Figure 5.3 Homology Model of CHA. Analysis of CHA protein domains with a putative model of CHA based on Glycerophoshphodiesterase (GpdQ) from *Enterobacter aerogenes* (Hadler et al., 2008) using SWISS-Model "Automated Project Mode".
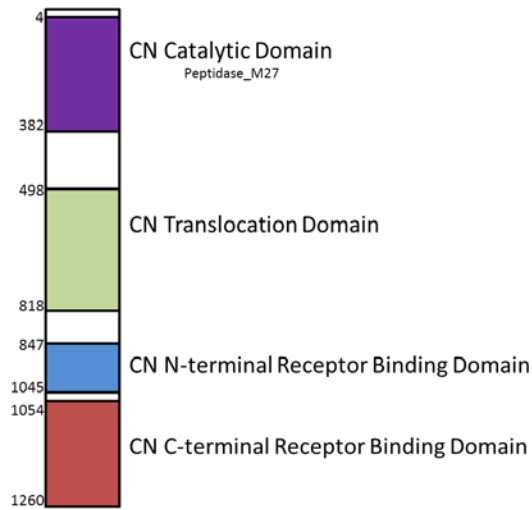
Botulinum neurotoxin type B complexed with synaptotagamin-II ectodomain PDB: 2np0.1.A
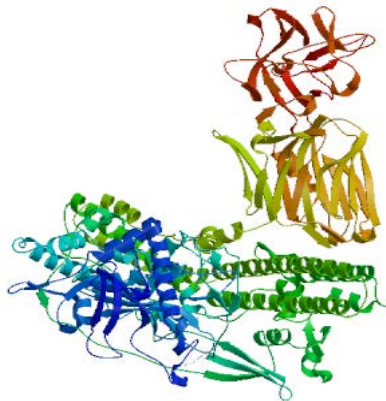
CMP1 model based on PDB: 2np0.1.A
Sequence Similarity: 31.57%
GMQE score: 0.68

Figure 5.4 Homology Model of CMP1. Analysis of CMP1 protein domains with a putative model of Cmp1 based on Crystal structure of the Botulinum neurotoxin type B complexed with synaptotagamin-II ectodomain (Chai et al., 2006) using SWISS-Model "Automated Project Mode".
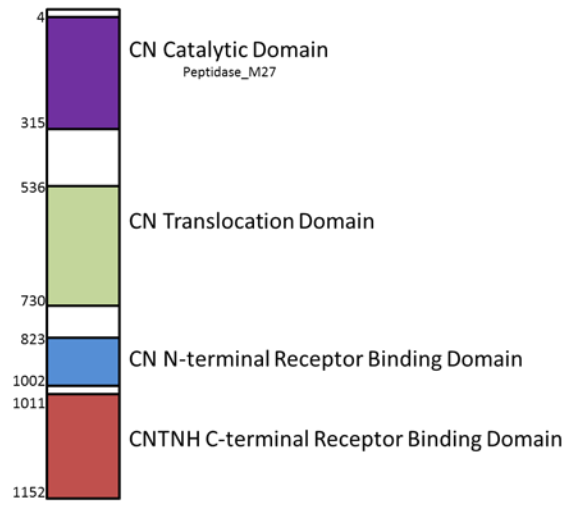
crystal structure of BoNT/Ai in
complex with NTNHA
PDB: 3v0a.1.B

CMP1 model based on PDB: 3v0a.1.B
Sequence Similarity: 43.19%
GMQE score: 0.74

Figure 5.5 Homology Model of CNTNH. Analysis of CNTNH protein domains with a
putative model of CNTNH based on crystal structure of BoNT/Ai in complex with
NTNHA (Gu et al., 2012) using SWISS-Model "Automated Project Mode".

Figure 5.6: Expression of Cmp operon rBT constructs in *Bacillus* expression vector. Only CMP1 showed expression in 4Q7 cells. CNTNH and CNTNH-HA expression in 4Q7 cells could not be tested as the antibody synthesized against CNTNH did not work.
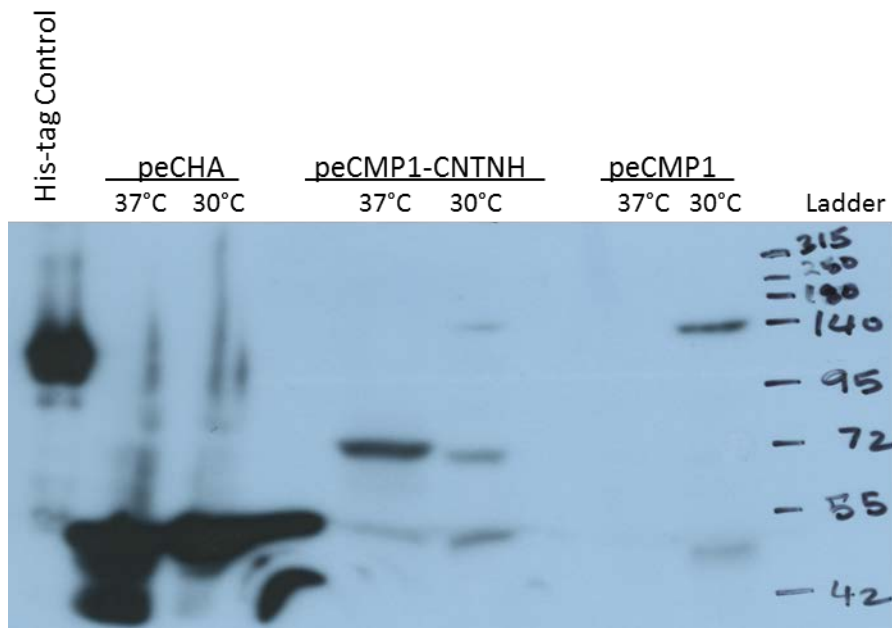
Figure 5.7: Expression of Cmp operon rEC constructs in *E. coli* expression vector.
Immunoblot using anti-6x-His-Tag antibody show all three peCHA, pe CMP1-CNTNH
and peCMP1, were expressed at 24 h. All the three constructs show more stable protein at
30°C than at 37°C. Hence the bioassays were done with cultures grown at 30°C.
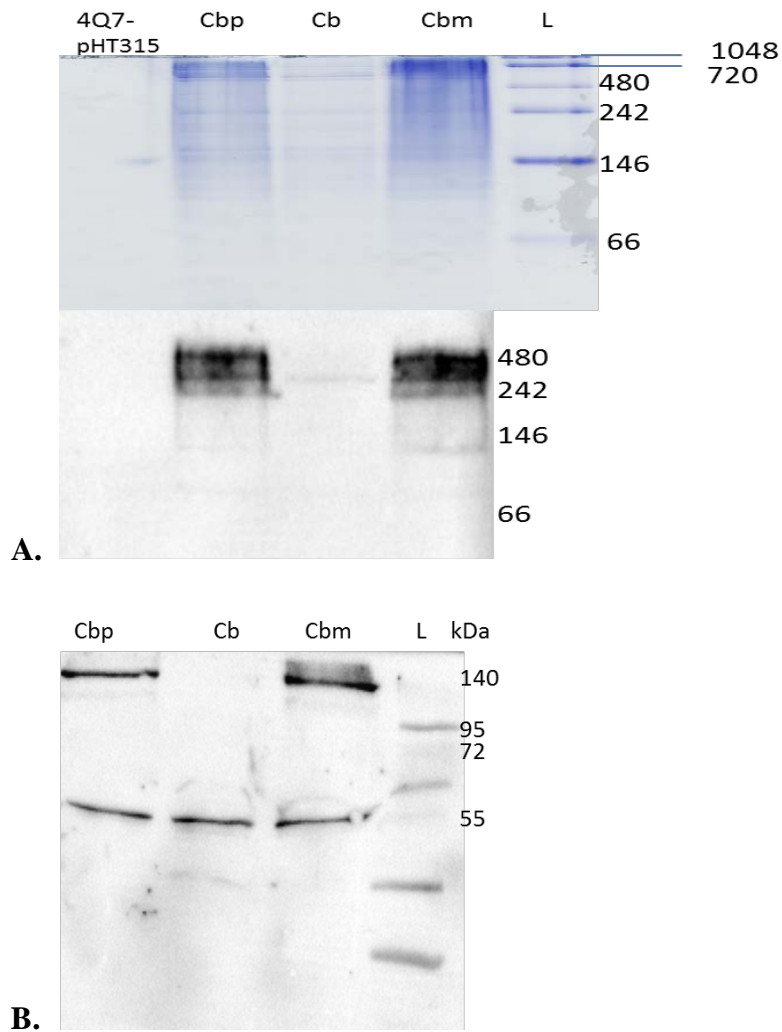
Figure 5.8: Expression of Cmp operon in Cbm, Cb and Cbp whole cultures using Anti-CMP1 antibody.

A. **Native PAGE blot shows the Cmp operon forms a complex** in Cbm and Cbp whole cultures (both are mosquitocidal and both have the pClosMP plasmid) but not in Cb which is not mosquitocidal.

B. **SDS-PAGE Immunoblot using anti-CMP1 antibody** show CMP1 is expressed as a 140 kDa protein in Cbm and Cbp. The CMP1 antibody cross-reacts with a non specific 55kDa clostridial protein..

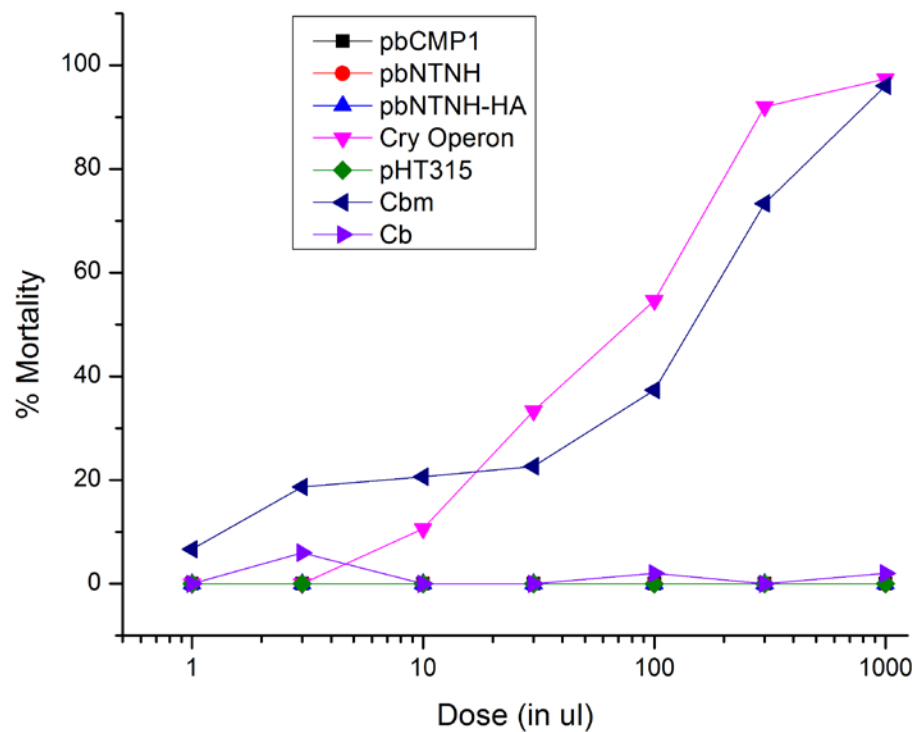**Figure 5.9: Cmp operon shows no toxicity to *Aedes* mosquitoes.**
Cbm cultures show high toxicity to *Ae aegypti*
Cry operon is toxic to *Aedes* mosquitoes.
pbCMP1, pbCNTNH and pbCNTNH-CHA all show no toxicity.
However, Cb and the rBT_pHT315 were nontoxic.
Potentially the use of a stronger *Bacillus thuringiensis* promoter, like Cyt1A, in case of
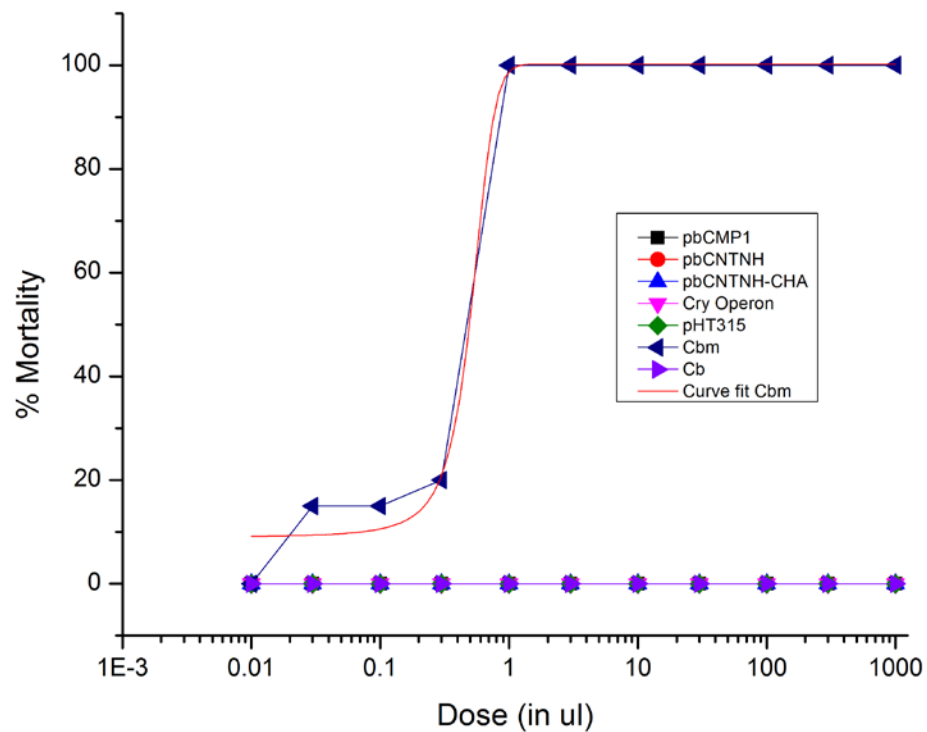rBt_CryO results in the production of high toxin levels observed here.

Figure 5.10: Cmp operon shows no toxicity to *Anopheles* mosquitoes.
Cbm cultures show high toxicity to *An. gambiae.*
Cry operon is not toxic to *Anopheles* mosquitoes.
pbCMP1, pbCNTNH and pbCNTNH-CHA all show no toxicity.
Also, Cb and the rBT_pHT315 were nontoxic.

Table 5.1. List of constructs used to assess toxicity of genes in the Cmp operon

Construct Name	Expression Vector used	Tag attached to protein	*Clostridium bifermentans* subsp. *malayasia* genes used
pbCMP1	pHT315	None	*CMP1*
pbCNTNH	pHT315	None	*CNTNH*
pbCNTNH-CHA	pHT315	None	*CNTNH, CHA*
peCMP1	pRSF™-1	6x-His	*CMP1*
peCMP1-NTNH	pRSF™-1	6x-His, S-tag	*CMP1, CNTNH*
peCHA	pET™-1	6x-His	*CHA*
pET-CHA-Stag	pET™-1	S-tag	*CHA*
pRSF-CMP1	pRSF™-1	None	*CMP1*

All "pb" constructs were in pHT315 using the Cyt1A promoter from *B. thuringiensis israelensis*

All "pe" constructs were in Duet vectors using the IPTG inducible T7 promoter

Table 5.2. List of primers used to create recombinant *E. coli* constructs

#	Primer Description	Sequence
1	pET-CHA-Stag-Forward	TT<u>CATATG</u>TCATTAAAAATCTACAATTCTAAA
2	pET-CHA-Stag-Reverse	TA<u>CTCGAG</u>CTTATATCAATAGTTTTAATTA
3	pRSF-CMP1-Forward	AAT<u>CCATGG</u>TACAAATAAGAGTTTTTAAT
4	pRSF-CMP1-Reverse	TA<u>GGATCC</u>TATTCCTTCCATCCTTCA

Table 5.3. Homology of the Cmp Operon genes

Cbm gene	Homology	Max. Identity	Model Created	% Identity to the model
HA	partial HA gene in *P. dendritiformis*	45%	Phosphohydrolase	16.88%
CMP	botulinum neurotoxin B	32%	Botulinum Neurotoxin B	31.57%
OXRF3	toxin complex component ORFX3 in *P. dendritiformis*	67%	None	
ORFX2	toxin complex component ORFX2 in *P. dendritiformis*	45%	SEX4 glucan phosphatase	21.28%
ORFX1	toxin complex component ORFX1	48%	Transcription Initiation Factor IIA	8.33%
NTNH	NTNH protein	42%	NTNH	42.41%
P47	partial P47 gene in *P. dendritiformis*	45%	Lumazine protein	26%

Table 5.4. Bioassay data against *Aedes aegypti* and *Anopheles gambiae* used in Figures 5.9. and 5.10. respectively. pbCMP1, pbCNTNH, pbCNHTNH-CHA, Cry operon and pHT315 cultures were grown for 24h at 30°C with shaking. Cbm and Cb cultures were grown in anaerobic jar for 18h at 30°C with no shaking. 25 third-instar *Aedes* larvae and third-to-fourth instar *Anopheles* larvae were used and put in 200ml water in plastic cups for bioassay. %Mortality was recorded after 24 hr.

% Mortality (average of 2-3 replicates) against *Aedes aegypti*							
Dose (in µl)	pbCMP1	pbCNTNH	pbCNTNH-CHA	Cb	pHTt315	Cbm	Cry Operon
1	0	0	0	0	0	6.66667	0
3	0	0	0	0	0	18.66667	0
10	0	0	0	0	0	20.66667	10.66667
30	0	0	0	0	0	22.66667	33.33333
100	0	0	0	0	0	37.33333	54.66667
300	0	0	0	0	0	73.33333	92
1000	0	0	0	0	0	96	97.33333
% Mortality (average of 2-3 replicates) against *Anopheles gambiae*							
Dose (in µl)	pbCMP1	pbCNTNH	pbCNTNH-CHA	Cb	pHTt315	Cbm	Cry Operon
0.01						0	
0.03						15	
0.1						15	
0.3						20	
1	0	0	0	0	0	100	0
3	0	0	0	0	0	100	0
10	0	0	0	0	0	100	0
30	0	0	0	0	0	100	0
100	0	0	0	0	0	100	0
300	0	0	0	0	0	100	0
1000	0	0	0	0	0	100	0

Chapter 6

General conclusions

The objective of this dissertation is to identify the mosquitocidal activity of *Clostridium bifermentans* serovar *malaysia* (Cbm) on *Aedes* and *Anopheles* sp. The research focused on completely sequencing the Cbm genome and to investigate the mosquitocidal activity of two toxin operons in *Aedes aegypti* and *Anopheles gambiae*.

Cbm, an anaerobic bacterial strain, is highly toxic to *Anopheles*, with lower toxicity to *Culex* and *Aedes aegypti* (Lee and Seleena, 1990b; Thiery et al., 1992). The biocontrol of mosquitoes for the past decade has been dominated by *Bacillus thuringiensis* subsp. *israelensis* (Bti) toxins. Cbm toxicity against *A. stephensi* is about 10 times higher, but its toxicity against *Culex pipiens* or *Aedes aegypti* is 10 times lower, than that of Bti (Thiery et al., 1992b). Cbm cells show no toxicity to mammals, goldfish (Thiery et al., 1992) and various other non-target organisms (Yiallouros et al., 1994). These characteristics make Cbm a promising target for identification of a biological insecticide for management of mosquito populations. Early reports (Barloy et al., 1996, 1998) indicated that Cbm toxins had mosquitocidal activity but a subsequent report in 2001 (Juarez-Perez et al., 2001) indicated the Cbm toxins are not mosquitocidal.

Cbm toxicity is presumably due to toxic protein(s) that is susceptible to proteases released when the cell lyses (Charles et al., 1990; Nicolas et al., 1990). Proteins in the Cbm mosquitocidal activity have not been clearly determined as they aggregate into a complex that is unstable, being inactivated by physical filtration, sonication, or a cycle of

217

freezing-thawing (Nicolas et al., 1993). Thus it was not previously possible to purify

them, without loss of toxicity.

In the first objective of this study, I analyzed and compared the genomic

sequences of *C. bifermentans* (Cb)*, C. bifermentans subsp. paraiba* (Cbp) and *C.*

bifermentans subsp malaysia (Cbm). All three bacterial strains were sequenced because

clostridial genomes tend to vary a lot in size and numbers of genomes and plasmids and

also physiology and their environmental conditions. Cbm and Cbp show mosquitocidal

activity whereas Cb the type strain is non mosquitocidal. Genomic DNA of all the strains

was sequenced and assembled de novo into large scaffolds. I closed the gaps between the

scaffolds using bioinformatics and PCR and assembled the complete genome of Cbm

with just one gap remaining. The whole genome of Cbm was also annotated. I found that

the genomes of larvicidal Cbm and Cbp differed from Cb by about 250kb. This could

only be explained by presence of plasmids, even though earlier reports had not found the

presence of a plasmid in Cbm (Selena et al., 1994). By computation analysis and

mutagenesis, I showed that Cbm contains eight plasmids, the largest of which is a

plasmid of 109kb (pClosMP) contained previously implicated *Cry16A*, *Cry17A*, *Cbm17.1*

and *Cbm17.2* genes in an operon. This was corroborated with previous research in my lab

that showed this plasmid was lost in the Cbm-mutant created by Dr. Qureshi. pClosMP

plasmid also carried a second toxin operon (named Cmp operon, Clostridial

Mosquitocidal Protein operon) encoding for Botulinum like toxin gene, Non Toxic Non

Hemagglutinin (*NTNH*) gene, *OrfX1*, *OrfX2*, *OrfX, p47* genes and Hemagglutinin (*HA*)

gene. The pClosMP plasmid also contains a putative large operon of about eighteen genes

which show homology to transport genes in other virulent bacteria and may play an important role in Cbm toxicity. The seven other plasmids ranging from sizes 2kb to 35kb were also analyzed. Annotation analysis of these plasmids show that out of these seven plasmids, four encode for only mobile elements and three contain homology to virulence associated genes which have played important role in other virulent bacteria. Two plasmids out of the latter three contain replication proteins and hence are self-replicating. This data provides a rich groundwork for further work on these plasmids.

The second objective of this research focused on investigating if the two toxin operons on the pClosMP plasmid are involved in Cbm mosquitocidal activity. This chapter was published in its entirety in Journal of Applied and Environmental Microbiology in September 2014 issue (Qureshi et al., 2014). We showed that the four toxins encoded by the Cry operon, Cry16A, Cry17A, Cbm17.1, and Cbm17.2, are all required for toxicity, and these toxins collectively show remarkable selectivity for *Aedes* rather than *Anopheles* mosquitoes, even though Cbm is more toxic to *Anopheles* (Qureshi et al.,2014). We show that when the Cbm Cry operon was expressed in its entirety, high levels of larvicidal activity towards *Aedes aegypti* was observed, and the toxicity of the 4Q7 *Bacillus* strain expressing the Cry operon, rBT_CryO, is higher than that observed with wild type Cbm (Figure 3.4B). However, to our surprise this 4Q7 *Bacillus* strain expressing the Cry operon showed no toxicity to *An. gambiae* and *C. quinquefasciatus* larvae. Our observations differ from those of Barloy et al., (1996) since we observed much higher levels of toxicity of the Cry operon to *Aedes* mosquitoes. Further, we but did not observe any toxicity to *Anopheles* or *Culex* mosquitoes, nor was any of the truncated

Cry operon constructs active as previously observed (Barloy et al., 1996). Hence, we concluded that the toxins which target *Anopheles* are different from those expressed by the Cry operon.

Next I investigated if the second operon on the pClosMP plasmid, Cmp operon, was involved in Cbm mosquitocidal activity. Based on homology to botulinum toxins and their mode of action, I selected three genes, *CMP1, CNTNH,* and *CHA* that play an important role in botulinum neurotoxicity. These were cloned and expressed in *Bacillus* strain but this 4Q7 *Bacillus* strain expressing the individual three genes showed no toxicity to *Aedes aegypti, Anopheles gambiae* and *Culex quinquefasciatus* larvae. Hence, I concluded that individually these three genes are not active against mosquitoes. Express the native full length 16kb Cmp operon and CMP1-CNTNH-HA in *Bacillus* strain constructs will give a complete picture.

The final objective was to investigate whether Cry17A plays an important role in toxicity of Cry Operon against *Aedes aegypti*. I investigated Domain II loop exchanges as a rational approach to modifying specificity based on previously reported site-directed mutagenesis and segment swapping experiments (Abdullah et al., 2003), but was unsuccessful. Further work is needed to create the loop mutants for both the approaches.

Consequently, this research demonstrated that the Cbm has acquired mosquitocidal genes through plasmid acquisition. The almost complete genome sequence of Cbm consists of a 3,723,765 bp chromosome encoding 3,835 ORFs. The previously implicated larvicidal Cry genes are encoded on a 109-kbp plasmid, containing 111ORFs. Additional virulence-related factors on the plasmids and chromosome could be identified

for further work. At least two toxin operons are present in Cbm. It seems Cry toxin operon is expressed as a 180Kda toxin complex as individual genes/proteins do not show mosquitocidal activity. The Cry toxin operon while active against *Aedes aegypti* are not active/appear to have negligible activity against *Anopheles gambiae* and *Culex quinquefasciatus* mosquitoes. Toxicity of Cbm towards *Anopheles gambiae* is still unknown and I hypothesize that it probably resides in the CMP locus which is expressed as a 480Kda toxin complex in the Cbm culture.

References

Abdullah, M. A., 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. *AEM*, **69**(9): 5343–5353

Barloy, F., Delécluse, A., Nicolas, L. and Lecadet, M. M. (1996). Cloning and expression of the first anaerobic toxin gene from *Clostridium bifermentans* serovar *malaysia*, encoding a new mosquitocidal protein with homologies to *Bacillus thuringiensis* delta - endotoxins. *J. Bacteriol.* **178**: 3099–3105.

Barloy, F., Lecadet, M. M. and Delécluse, A. (1998). Cloning and sequencing of three new putative toxin genes from *Clostridium bifermentans* CH18. *Gene* **211:** 293–299.

Charles, J. -F., L. Nicolas, M. Sebald, and H. de Barjac. (1990). *Clostridium bifermentans* serovar *malaysia*: sporulation, biogenesis of inclusion bodies and larvicidal effect on mosquito. *Res. Microbiol.* **141:** 721–733.

Ibrahim, M. A., Griko, N. B. and Bulla, Jr. L.A. (2013). The Cry4B toxin of *Bacillus thuringiensis* subsp. *israelensis* kills Permethrin-resistant *Anopheles gambiae*, the principal vector of malaria. *Exp Biol Med (Maywood).* **238**(4):350-9.

Juárez-Pérez, V. and Delécluse, A. (2001). The Cry toxins and the putative hemolysins of *Clostridium bifermentans* ser. *malaysia* are not involved in mosquitocidal activity. *J Invertebr Pathol.* **78(1):** 57-8.

Lee, H. P. and Seleena, P.(1990). Isolation of indigenous larvicidal microbial control agents of mosquitoes: The Malaysian experience. S*outheast Asian J. Trop. Med. Public Health.* **21**: 281–287.

Nicolas, L., Charles, J.-F., and de BarjaC, H. (1993). *Clostridium bifermentans* serovar *malaysia*: characterization of putative mosquito larvicidal proteins. *FEMS Microbiol. Lett.* **113:** 23–28.

Nicolas, L., Hamon, S., Frachon, E., Sebald, M. and de BarjaC, H. (1990). Partial inactivation of the mosquitocidal activity of *Clostridium bifermentans* serovar *malaysia* by extracellular proteinases. *Appl. Microbiol. Biotechnol.* **34:** 36–41.

Qureshi, N., Chawla, S., Likitvivatanavong, S., Lee, H. L., and Gill, S.S. (2014). The cry toxin operon of *Clostridium bifermentans* subsp. *malaysia* is highly toxic to *Aedes* Larval Mosquitoes. *Appl Environ Microbiol*. **80**(18):5689-97.

Seleena, P. and Lee, H. P. (1994). Absence of plasmid in mosquitocidal *Clostridium bifermentans* serovar *malaysia*. *Southeast Asian J. Trop. Med. Public Health* **25(2):** 394-396

Thie´ry, I., Hamon, S. V., Dumanoir, C. and de BarjaC, H. (1992a). Vertebrate safety of *Clostridium bifermentans* serovar *malaysia*, a new larvicidal agent for vector control. *J. Econ. Entomol.* **85:** 1618–1623.

Thie´ry, I., Hamon, S., Gaven, B. and de BarjaC, H. (1992b). Host range of *Clostridium bifermentans* serovar *malaysia*, a mosquitocidal anaerobic bacterium. *J. Am. Mosq. Control Assoc.* **8:**272–277

Yiallouros, M., Storch, V., Thie´ry, I. and Becker, N. (1994). Efficacy of *Clostridium bifermentans* serovar *malaysia* on target and nontarget organisms. *J. Am. Mosq. Control Assoc.* **10:**51–55.