# Isolation and characterization of parasporal anticancer protein from *Bacillus thuringiensis* strains

**Ph.D. Thesis**

**Md. Mahmud Hasan**


**Department of Microbiology**
**Faculty of Biological Science**
**University of Dhaka**
**Dhaka, Bangladesh**

**July 2024**

# Isolation and characterization of parasporal anticancer protein from *Bacillus thuringiensis* strains

**A DISSERTATION SUBMITTED FOR THE DEGREE OF**

**DOCTOR OF PHILOSOPHY**

**SUBMITTED BY:**

**MD. MAHMUD HASAN**

**REGISTRATION NO: 104**

**SESSION: 2019-2020**


**Department of Microbiology**
**Faculty of Biological Science**
**University of Dhaka**
**Dhaka-1000, Bangladesh**

**July 2024**

## Dedicated to...

My beloved parents, my wife Sarah Farhana Moon, and my daughters, Sidratul Muntaha and Sirajam Muneera, who cherished my life with their love and affection.

## *Quotation…*

"My Lord! Grant me an honorable entrance and an honorable exit, and give me a supporting authority from Yourself."

Surah Al-Isra, verse 80, Holy Quran

# CERTIFICATE OF APPROVAL

| | |
|---|---|
| **Name** | Md. Mahmud Hasan |
| **Registration number and Session** | Reg. no. 104, 2019-20 |
| **Department** | Microbiology, University of Dhaka |
| **Degree** | Doctor of Philosophy |
| **Title of the Thesis** | Isolation and characterization of parasporal anticancer protein from *Bacillus thuringiensis* strains |

We approve the submission of his PhD Dissertation.

Date of Signature

.....SNKhan..... July 14, 2024

**Professor Dr. Shakila Nargis Khan**
Supervisor
Department of Microbiology.
University of Dhaka
Dhaka-1000, Bangladesh.

..... July 14, 2024

**Professor Dr. Md. Mozammel Hoq**
Co-Supervisor
Department of Microbiology.
University of Dhaka
Dhaka-1000, Bangladesh.

..... July 14, 2024

**Professor Dr. Anowara Begum**
Co-Supervisor
Department of Microbiology
University of Dhaka
Dhaka-1000, Bangladesh.

# ABSTRACT

Extensive research is ongoing to unravel the mystery of the diversified role of Cry proteins from *Bacillus thuringiensis* (Bt). Anticancer Cry proteins are named parasporins and have the selectivity to kill tissue-specific cancer cells without harming normal cells. These characteristics of parasporins have opened a new horizon of its anticancer activities. This study sought for investigation into new parasporal anticancer proteins from 98 nonhemolytic and 23 non-insecticidal hemolytic Bt strains from 298 indigenous Bt stock. Six sets of primers were designed from the conserved *δ-endotoxin*, *CryBP1*, and *Etx-Mtx* regions of *ps1- ps5* and were explored to detect the presence of respective genes through PCR amplification. A total of 27 PCR amplicons were obtained from 121 Bt strains. Primers targeting the *ps1* and *ps3 (CryBP1)* genes amplified desired fragments from the plasmid of 8 Bt strains. Besides, *ps1, ps3* and *ps4* primers amplified 19 unintended amplicons of various sizes. 25 PCR amplicons were sequenced from both anticipated and unintended amplifications. These generated no expected parasporin or cry genes from the DNA sequencing data, instead 11 plasmid-mediated genes were found, mostly of drug/metabolite transporters, Tn3 transposases, and DNA/RNA manipulators.

Alternatively, extraction and purification of parasporal proteins followed by SDS-PAGE were performed from 98 nonhemolytic Bt strains. Of them, seven strains; MyIa2, FHSc4, KkSc1, RaSd1, MyIb1, MyIa1, and SgSp2 revealed some prominent parasporal protein bands with varied molecular masses (17-78 kDa). Among them, proteinase K-digested putative parasporal proteins from 3 strains, MyIa2, FHSc4, and KkSc1, were revealed as cytotoxic to the HeLa cell line. Phase-contrast microscopy also supported the presence of parasporal inclusion bodies from the same three strains. Interestingly, as proteinase K digested parasporal protein did not produce any visible protein band in the SDS gel suggesting that they belong to non-Cry or novel/ undiscovered categories of parasporal proteins which deserve more critical studies. Consequently, the research was covered to investigating the anticancer Cry protein in 23 hemolytic non-insecticidal Bt strains of which five strains produced putative parasporal protein bands on the SDS-PAGE. Bt strains Soil-46 and 28S were found to produce four putative parasporal proteins of identical molecular mass (101, 86, 76, and 28 kDa), whereas BD59S had three (101, 86, and 76 kDa). Besides, a 103 kDa protein band was found from Dsh4 and 45L strains. Approximately 64, 51, 45, 41, and 32 kDa fragments were generated from the putative parasporal protein extracted from 28S and Soil-46 upon tryptic digestion. BD59S had 64, 51, and 41 kDa protein fragments, while Dsh4 and 45L had 45 kDa fragments upon trypsin digestion. On the other hand, parasporal proteins from all five strains formed a 46 kDa identical protein fragment when digested by proteinase K. Then, untreated, trypsin-treated, and proteinase K-treated protein samples from each strain were tested for cytotoxicity on HeLa cancer cell lines. Untreated protein samples from 28S, Soil-46, 45L, and Dsh4 exhibited modest cytotoxicity on the HeLa cell line, while BD59S did not. The trypsin-digested proteins of BD59S exhibited significant cytotoxicity, whereas 28S, Soil-46, 45L, and Dsh4 showed moderate towards the HeLa cell lines. Proteinase K digested proteins of all five strains showed potent

cytotoxicity on the HeLa cell line. In contrast, the same protein samples at the same concentration didn't show any cytotoxicity to the normal cell, Vero cell line. In addition, the solubilized parasporal proteins of these Bt strains showed cytotoxicity to brine shrimp nauplii with an $LC_{50}$ value of 114.19 µg/mL. Furthermore, they were nonhemolytic to sheep red blood cells and non-insecticidal to the fruit fly larvae (*Bactrocera cucurbitae*). All of these findings meet the requirements of being parasporins. Besides, as the untreated parasporal protein from 4 strains Soil-46, 28S, 45L, and DSh4 showed little toxicity towards the HeLa cell line hence this attitude of parasporal proteins someway coordinated with S-layer protein. Light microscopy showed the presence of round/ atypical parasporal inclusions, while SEM showed that all inclusions were round in five strains. Morphology analysis from SEM indicates that the parasporal inclusions are consistent with the S-layer protein.

Whole genome sequencing (WGS) of four strains (BD59S, 28S, Soil-46, and 45L), three strains (BD59S, Soil-46, and 28S) were found belonging to Bt serovar *finitimus*, whereas 45L belonging to the serovar *konkukian*. Besides, bacterial WGS data analysis, the 101, 86 and 76 kDa SDS-PAGE observed protein bands were identified as 91 kDa, 86 kDa S-layer, and a three domain Cry-like 73.54 kDa protein, respectively. The 103 kDa protein band of 45L was identified as 103 or 93 kDa S-layer protein of these strains. As anticancer activity was identified in proteinase K and trypsin digested parasporal protein, hence they were compared with anticancer parasporin. The 91 kDa EA1, 86 kDa Sap, and 73.54 kDa Cry-like protein had 19% and 18% identity with PS3, respectively. Besides, the 93 kDa SLP had 16% with PS3, and the 103 kDa SLP had 19.84% and 14.22% similarities with PS3 and PS5, respectively.

Based on the results, parasporal anticancer proteins of five indigenous Bt strains namely *Bacillus thuringiensis* BD59S, Soil-46, 28S, 45L, and Dsh4 are suggested to be as parasporin-like S-layer proteins. This is the first report of S-layer proteins from *B. thuringiensis* having anticancer potential in the HeLa cancer cell line while non-cytotoxic to the Vero cell line. The cancer-cell killing nature of parasporal S-layer protein from Bt strains thus opens up further investigations to unveil their anticancer potentials to become a candidate for anticancer protein development.

# ACKNOWLEDGEMENTS

First of all, I want to attribute all commendation to the infinitely Almighty Allah for granting me the ability to do my doctoral study and successfully submit my thesis.

I want to extend heartfelt acknowledgment, profound admiration, and gratitude to my esteemed supervisor, Dr. Shakila Nargis Khan, Professor and Chairman of the Department of Microbiology at the University of Dhaka. Her unwavering support, uncompromising dedication, and thorough guidance and direction have been invaluable in monitoring and evaluating my PhD research.

I want to convey immense appreciation, genuine admiration, and gratitude to my respected mentor and Co-supervisor, Professor Dr. Md. Mozammel Hoq of the Department of Microbiology at the University of Dhaka, for providing me with invaluable advice and unwavering support for the whole of my doctoral study.

I want to express my sincere appreciation and extend heartfelt gratitude to my respected mentor and Co-supervisor, Dr. Anowara Begum, Professor in the Department of Microbiology at the University of Dhaka, for her insightful and valuable comments, as well as her consistent support and collaboration throughout my Ph.D. study.

I want to thank Professor Dr. Muhammad Manjurul Karim, and Professor Dr. Marufa Zerin Akhter in the Department of Microbiology at the University of Dhaka, for their timely recommendations, and unwavering support throughout my research for my Ph.D. thesis.

I express my gratitude to all the current and former members of the Fermentation and Enzyme Biotechnology Laboratory (Lab-215) during my time there, namely Dr. Mukitu Nahar, Dr. Md. Asaduzzaman Shishir, Mr. Shafayet Imtiaze Khan, Avirup Saha, Hasib Adnan, and Md. Mahmuduzzaman Mian for their significant scientific and technical assistance.

I express my gratitude to the whole personnel of this department for their genuine assistance during my research endeavor.

I would like to express my appreciation to Dr. Sheikh Ariful Hoque, Principal Scientist; Dr. Tania Hossain, Scientist at the Cell and Tissue Culture Laboratory, Centre for Advanced Research in Sciences (CARS); Md. Anowar Hosen, a scientist at the Scanning Electron Microscope Laboratory of CARS, University of Dhaka. The NSU Genome Research Institute (NGRI) of North South University and BCSIR to conduct whole genome sequencing.

I express my heartfelt gratitude to the National Institute of Biotechnology (NIB) for issuing me study leave, and to the Bangabandhu Science and Technology Fellowship Trust for granting me a PhD fellowship to support my Ph.D. studies.

Lastly, I would like to express my gratitude to my wife, Sarah Farhana Moon, for her unwavering patience and consistent support during my doctoral study, which played a crucial role in enabling me to complete my thesis. I want to extend my heartfelt appreciation to my parents, father-in-law, and mother-in-law for their unwavering encouragement and limitless support during my journey to get this degree.

The Author
July 14, 2024.

# TABLE OF CONTENTS

# LIST OF FIGURES

| Figure No. | Figure Caption | Page No. |
|---|---|---|

| Figure No. | Figure Caption | Page No. |
|---|---|---|

# LIST OF TABLES

# ABBREVIATIONS AND SYMBOLS

| | |
|---|---|
| % | Percentage |
| ~ | Tilde, here used to indicate a range |
| aa | Amino acid |
| BLAST | Basic local alignment search tool |
| BLASTn | Nucleotide BLAST |
| BLASTp | Protein BLAST |
| bp | Base pair |
| Bt | *Bacillus thuringiensis* |
| CDD | Conserved domain database |
| CDS | Coding DNA Sequence |
| ClustalW | Clustal alignment format without base/residue numbering |
| cm | Centimeter |
| Cry | Crystalline inclusion protein of *Bacillus thuringiensis* |
| Cyt | Cytolytic inclusion protein of *Bacillus thuringiensis* |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxy nucleotide triphosphate |
| DTT | Dithiothreitol |
| EA1 | Extractable antigen-1 |
| EBI | European bioinformatics institute |
| EDTA | Ethylene diamine tetra-acetic acid |
| *et al.* | And others |
| Et-Br | Ethidium bromide |
| Etx | Epsilon toxin |
| × g | g-force or relative centrifugal force (rcf) |
| g | gram |
| HeLa | HeLa is a cervical cancer cell line derived from the name of the patient, Henrietta Lacks |
| Ig | Immunoglobulin |
| *in silico* | Silicon in computer chips / in or on a computer |
| kb | Kilobase |
| kDa | Kilodalton |
| Kg | Kilogram |
| L | Liter |
| LB | Luria bertani |
| M | Molar |

| | |
|---|---|
| MAFFT | Multiple alignment using fast fourier transform |
| MEGA | Molecular evolutionary genetics analysis |
| mg | Milligram |
| mL | Milliliter |
| mm | Millimeter |
| mM | Millimole |
| Mtx | Mosquitocidal toxins |
| MUSCLE | Multiple sequence comparison by log- expectation |
| MW | Molecular weight |
| NA | Nutrient agar |
| NCBI | National centre for biotechnology information |
| ng | Nanogram |
| NGS | Next generation sequencing |
| nm | Nanometer |
| OD | Optical density |
| PAGE | Polyacrylamide gel electrophoresis |
| PBS | Phosphate Buffered Saline |
| PCR | Polymerase chain reaction |
| PDB | Protein data bank |
| PGAP | Prokaryotic genome annotation pipeline |
| pH | Negative logarithm of hydrogen ion concentration |
| PS | Parasporin |
| RAST | Rapid annotations using subsystems technology |
| RNA | Ribonucleic acid |
| rpm | Rotation per minute |
| rRNA | Ribosomal RNA |
| RT | Room temperature |
| Sap | Surface array protein |
| SD | Standard deviation |
| SDS | Sodium dodecyl sulfate |
| sec | Second |
| SEM | Scanning electron microscope |
| S-layer | Surface layer |
| SLH | Surface layer homology |
| SLP | Surface layer protein |
| SMART | Simple modular architecture research tool |
| *spp.* | Species |

| | |
|---|---|
| *Taq* | *Thermus aquaticus* |
| TBE | Tris-borate EDTA |
| TE | Tris EDTA buffer |
| TEMED | Tetramethylethylenediamine |
| UniProt | Universal protein resource |
| UV | Ultraviolet |
| v/v | Volume per volume e.g. (mL/mL) |
| Vero | Kidney epithetical cell line derived from an African Green Monkey. Vero gets its name from a derivation of green kidney, "Verda Reno" |
| Vip | Vegetative insecticidal protein |
| vol. | Volume |
| w/v | Weight per volume e.g. (g/mL) |
| WGS | Whole genome sequencing |
| α | Alpha |
| β | Beta |
| γ | Gamma |
| δ | Delta |
| µ | Micro |
| µg | Microgram |
| µL | Microliter |
| µm | Micrometer |
| 3D | Three dimensional |

# ABBREVIATED NAMES OF AMINO ACIDS

G - Glycine

V - Valine

L - Leucine

I - Isoleucine

F - Phenylalanine

P - Proline

Y - Tyrosine

W - Tryptophan

S - Serine

T - Threonine

A - Alanine

M - Methionine

N - Asparagine

Q - Glutamine

D - Aspartate

E - Glutamate

K - Lysine

R - Arginine

C - Cysteine

H - Histidine

# CHAPTER 1:
# INTRODUCTION

# CHAPTER 1: INTRODUCTION

*Bacillus thuringiensis* (Bt) is a naturally occurring, Gram-positive, rod-shaped, endospore-forming soil-borne bacterium that could be either a facultative or obligate aerobe (Drobniewski, 1993), commonly used as a biological alternative to a pesticide. Bt was first isolated in Japan by Shigetane Ishiwatari (Ishiwatari, 1901) as a pathogen of the sotto disease of silkmoth (*Bombyx mori*) larvae and named *Bacillus sotto*. In 1911, Ernst Berliner detected Bt in Germany (Berliner, 1915). He extracted it from a sick Mediterranean flour moth (*Ephestia kuehniella*) in the German state of Thuringia and gave it the name *Bacillus thuringiensis*. Bt belongs to the *B. cereus* group, which comprises five more species: *B. cereus, B. anthracis, B. weihenstephanensis, B. pseudomycoides,* and *B. mycoides*. It has a high degree of similarity in both biological and genetic characteristics to *B. cereus* and *B. anthracis* (Guttmann & Ellar, 2000).

Bt is well known for its ability to produce parasporal proteinaceous crystalline inclusions (Cry) during sporulation. *B. thuringiensis* has around 150 distinct subspecies. Each of the subspecies can synthesize one or more parasporal crystalline toxins that belong to either pore-forming toxins (PFT) or δ- endotoxins. Certain strains of *B. thuringiensis* also synthesize S-layer proteins as round parasporal inclusions that are toxic to certain pests, mosquito larvae, and cancer cell lines (Lormendez *et al.,* 2019; Rubio *et al.,* 2017). Cry proteins have unique and specific poisonous actions against selective insect species. Due to their natural propensity to form inclusion bodies, the Cry toxins of Bt may be easily extracted. These toxins have garnered significant interest as biological control agents, considering their potential in managing agricultural insect pests and insect vectors of human illnesses. However, Bt toxin poses little harm to the environment, people, or non-target creatures because of its limited ability to affect specific hosts and its effectiveness in managing diverse insect pests (Bechtel & Bulla, 1982; Bravo *et al.*, 2011, 2013; Drobniewski, 1993; Walters *et al.*, 2010). δ-endotoxins provide an acceptable replacement for synthetic pesticides in the management of insect pests, making them a preferred choice among other toxins (Sung *et al.*, 2021).

In 1999, for the first time in Bt research, 3 non-insecticidal parasporal Cry proteins from non-hemolytic Bt strains were discovered by Japanese scientists Mizuki and his team that had selective cytotoxicity against human leukemia cells (Mizuki *et al.,* 2000). These findings led to the screening of non-insecticidal *B. thuringiensis* strains globally to identify parasporal anticancer Cry proteins. Parasporal Cry proteins that are nonhemolytic, and noncytotoxic to normal cell lines but are capable of killing human cancer cells are called parasporins. The cancer cell-killing mechanisms of action of parasporins are two types: the three-domain Cry toxin type (PS1, PS3, PS6), cell death by apoptosis via receptor-mediated signaling; and the β-pore-forming toxin (*β*-PFT) type (PS2, PS4, PS5), necrotic cell death by oligomerizing and pore-forming toxin via multiple receptors. The abundance, diversity, and selectivity of Cry proteins in Bt have made them potential candidates for cancer treatment (Wong, 2010), which will have great implications in medical research for the development of potential therapeutics for certain cancers or as diagnostic agents.

In addition to parasporal Cry protein, *B. thuringiensis* also produced another type of parasporal inclusion protein, called surface layer protein (SLP). In 2001, for the first time, a 100 kDa surface layer protein (SLP) was identified to produce atypical oval-shaped proteinaceous parasporal inclusion bodies in the *B. thuringiensis* strain CTC (Sun *et al.,* 2001). Later, it was proved by several reports that SLPs are a new group of parasporal inclusions and about 25% of parasporal inclusions of *B. thuringiensis* could possess SLP (Guo *et al.,* 2008). It was experimentally proved that *B. thuringiensis* strains, CTC, B22, I13, H67, and BMB1152 produced two different types of parasporal SLPs, the surface array protein (Sap), and the extractable antigen-1 (EA1), (Zhou *et al.*, 2011). Experimental evidence confirms that the surface array protein coding gene *sap* and the extractable antigen-1 protein coding gene *eag* are found together in 35% of the analyzed *B. thuringiensis* strains, whereas they are lacking in 65% of the strains. Most of the parasporal SLPs of *B. thuringiensis* were found as non-toxic and the function of the SLP coding gene is still unclear (Rubio *et al.,* 2017). The contribution of S-layer proteins of *Lysinibacillus sphaericus* as a mosquitocidal agent (Allievi *et al.,* 2014; Thanabalu *et al.,* 1991) is well known. Some recent studies showed that SLPs of *B. thuringiensis* had selective toxicity to

coleopteran beetles (Peña *et al.*, 2006), cattle ticks (Lormendez *et al.,* 2019), and human cancer cells (Rubio *et al.,* 2017).

Cancer is a large group of diseases that can start in almost any organ or tissue of the body when abnormal cells grow uncontrollably, go beyond their usual boundaries to invade adjoining parts of the body and/or spread to other organs. An estimated 19.3 million people were diagnosed with cancer, and nearly 10 million lost their lives in 2020 (Sung *et al.*, 2021). This accounts for roughly one in six fatalities. Cancer can develop almost anywhere in the cells of the body, and once cancerous, those cells may multiply uncontrollably and migrate to other organs of the body. There are over 200 different types of cancer (https://www.cancerresearchuk.org/what-is-cancer/how-cancer-starts/types-of-cancer), and they are typically designated by the organs or tissues where they first appear; for instance, lung cancer normally develops in the lung and brain cancer in the brain. Most cases of cancer are found in the breast, lungs, colon, cervix, ovary, kidney, white blood cells, brain, and prostate. Cancer can be classified into five main categories based on the type of cell it originates from. These categories include carcinoma (originating in the skin or tissues that line or cover internal organs), sarcoma (originating in connective or supportive tissues like bone, cartilage, fat, muscle, or blood vessels), leukemia (starting in the tissues responsible for producing blood cells, such as the bone marrow), lymphoma and myeloma (originating in immune system cells), and brain and spinal cord cancers (central nervous system cancers).

It has been demonstrated that certain proteins and peptides could selectively eliminate specific cancer cells. Parasporal proteins parasporin of *B. thuringiensis* also exhibit selective abilities to eliminate certain cancer cells. Besides, S-layer proteins have also been found to show anticancer activity against some cancer cell lines, and they have also been synthesized as parasporal inclusions in some *B. thuringiensis* strains. The parasporal proteins of *B. thuringiensis* possess targeted biocidal, cytocidal and anticancer properties, and the ease of their production, processing, and extraction on a laboratory scale made them popular worldwide. Thus, the parasporal proteins that have the ability to destroy cancer cells in a tissue-specific manner could be

investigated as potential new therapeutic candidates for cancer. With the concern of developing anticancer therapeutic, this study was, therefore, started with the screening of the Bt collection library of the Enzyme and Fermentation Biotechnology Laboratory of the Department of Microbiology, University of Dhaka, for the presence of potential anticancer parasporins at the genomic and proteomics levels. The results of this study would be expected to provide a clear understanding of the presence of parasporal anticancer proteins in indigenous Bt strains. Based on the successful research outcome, the cytotoxic effect of the parasporal proteins of indigenous Bt strains could be considered as potential candidates to develop targeted anticancer protein therapeutics in future.

# CHAPTER 2:

# REVIEW OF LITERATURES AND OBJECTIVES

# CHAPTER 2: REVIEW OF LITERATURES AND OBJECTIVES

## 2.1 *Bacillus thuringiensis*

*Bacillus* are a group of rods-like, Gram-positive bacteria that can produce endospores. Some of these bacteria can survive with or without oxygen, while others need oxygen to survive (Drobniewski, 1993). *Bacillus thuringiensis* is belonging *Bacillus cereus* group of bacteria (Rasko *et al.*, 2005), and is phenotypically and genotypically identical to *B. cereus* and *B. anthracis* (Guttmann & Ellar, 2000). *B. thuringiensis* is a soil bacterium that is motile (through peritrichous flagella), non-capsulated, indole, Voges Proskauer (VP), citrate, gelatin, as well as starch hydrolysis positive, but Methyl Red (MR) negative. *B. thuringiensis* sporulation is biochemically identical to other spore-forming bacteria, and its spores are comparable in appearance and content. *B. thuringiensis* was found to be different from other species of the *Bacillus cereus* group in that it formed parasporal inclusion bodies close to its spores during sporulation (Andrews *et al.,* 1987; Donald *et al.,* 1982; Gleave, *et al.,* 1992; Stahly *et al.,* 2006), which were poisonous to many insects (Palma *et al.,* 2014) and cancer cell lines (Gonzalez *et al.,* 2011; Jung *et al.,* 2007; Lenina *et al.,* 2014; Mizuki *et al.,* 1999; Mizuki *et al.,* 2000; Ohba *et al.,* 2009; Poornima *et al.,* 2010; Uemori *et al.,* 2007). Over 200 distinct cry proteins were produced by the thousands of various Bt strains. These parasporal inclusions become visible under a microscope and are insoluble in normal water (Fig 2.1). However, after being treated at 37 ºC for 1 hour in a 50 mM $Na_2CO3$ solution with 10 mM dithiothreitol (DTT), they became water-soluble (Brasseur *et al.*, 2015).

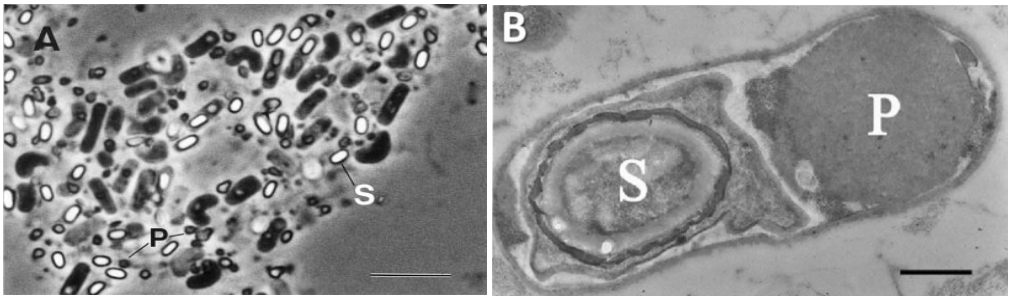



**Figure 2.1:** Microscopic observation of parasporal inclusion proteins of *B. thuringiensis.* (A) Micrograph displaying the phase-contrast of a culture of *B. thuringiensis* strain A1547 (formerly 90-F-45-14) undergoing sporulation (Kim *et al.,* 2000). (B) Transmission electron micrograph of a fully developed sporangium of *B. thuringiensis* BFR1 (Cry9Ec1) that contained a spore and a parasporal inclusion. S represents spore, P represents parasporal inclusion, Bar = 500 nm, (Wasano *et al.*, 2005).

## 2.2 The history of *B. thuringiensis* research and its recent advances

Japanese biologist Shigetane Ishiwatari investigated silkworm larvae sickness (sudden-collapse disease). The infection killed millions of silkworm larvae in Japan. Ishiwata isolated and named the bacterium sotto from dead silkworm larvae (Ishiwatari, 1901). Ten years later, in 1911, German scientist Ernst Berliner found a dead Mediterranean flour moth larva, from which he isolated a similar bacterium and named it *Bacillus thuringiensis*. Berliner observed in 1915 that the bacteria formed inclusion bodies or "Restkorper" along with the endospore (Berliner, 1915), but the activity of this crystal had not yet been identified at that time. In 1938, the first commercial insecticide, Sporeine, made from a spore-crystal complex, was marketed in France to combat wheat moths. Steinhaus showed in 1951 that Bt could be used to get rid of alfalfa caterpillars, it wasn't well-known as a pesticide (Steinhaus, 1951). With his discovery, a new era of research on Bt began, and later studies showed that the bipyramidal parasporal inclusion of Bt caused other caterpillars to die quickly. In 1953, Thomas Angus revealed that the inclusion bodies of *B. thuringiensis* were responsible for its insecticidal function (Angus, 1954). Most parasporal inclusions have a fine crystalline structure, and (Hannay, 1961) was the first to notice this. The development of Bt as an insecticide and its ability to compete with chemical pesticides was sped up by the fact that it could be fermented on a large scale. Parasporal crystal was found as the most effective insecticide against lepidopteran (moth) insects (Hannay & Fitz-James, 1955). Because of this finding, more attention has been paid to Bt's crystal structure, its biology, and how it works in general. Bt was commercialized in the US in 1958. Since then, Bt has been sprayed on crops and forests, and control disease vectors. The United States Environmental Protection Agency (EPA) registered Bt as a pesticide in 1961. In 1967, the term "parasporal crystalline protein" was coined by Heimpel to describe Bt, which crystallizes within the cell and includes toxic components (Heimpel, 1967). Commercial *B. thuringiensis* subsp. kurstaki HD1 products (Dipel and Thurid.de) were widely utilized to eradicate lepidopteran pests from agricultural and forested areas during the 1970s. After this achievement, 13 Bt subspecies were identified, but none were active against insects outside the order Lepidoptera. Howard Dulmage found the *B. thuringiensis* subsp. kurstaki HD1 strain to fight lepidopteran pests. This strain was found to be safe to use around bees, predatory insects, parasitoids, and other vertebrates (Margaritis & Dulmage, 1987).

In the 1976s, Prasad and Shethna discovered that a highly pure protein with anticancer properties, derived from the proteinaceous crystal of *B. thuringiensis*, effectively hinders the development of Yoshida ascites sarcoma (YAS) cells, both in *in vivo* and *in*

*vitro*. They discovered that the anticancer Cry protein changed the permeability of tumor cells, and suggests that it acts primarily on the tumor cell membrane (Prasad and Shethna, 1976).

(Goldberg & Margalit, 1977) identified *B. thuringiensis* subspecies israelensis as the first subspecies of this bacterium that was lethal to dipteran (fly) species. Its spherical inclusion bodies were very poisonous to the larvae of many different kinds of mosquitoes, flies, and chironomid midges of order Diptera. It performed an important role in the reduction of river blindness disease in West Africa caused by a nematode and transmitted by blackflies (Guillet *et al.,* 1990). When scientists and environmentalists realized that synthetic pesticides were harming the environment and when insects developed resistance to them, Bt usage skyrocketed in the 1980s. Bt is an organic substance that selectively targets certain insects and does not have a long-lasting presence in the environment. This prompted both public and business sectors to begin supporting Bt research. In 1987, A pathotype of Bt having toxicity to larvae and adults of coleopteran insects was discovered in Germany. It produced a thin rhomboidal parasporal crystal inclusion and named *B. thuringiensis* var. *tenebrionis* (Krieg *et al.,* 1987). It was developed as a bacterial insecticide for the control of Colorado potato beetle pests *(Leptinotarsa decemlineata).*

In 1988, non-insecticidal Bt strains were discovered. Out of 189 *B. thuringiensis* isolates that had parasporal inclusions (PIs), 36 % were toxic to various insect orders, while 64 % had no insecticidal activity. This indicated that in natural environments, the number of non-insecticidal PI-forming bacteria was greater than that of insecticidal ones (Ohba, 1996; Ohba & Aizawa, 1988). The United States EPA had registered a total of 182 Bt products by 1995. Currently, there are several variations of Bt. A significant number of them possess genes that encode distinct poisonous crystals inside their DNA. Due to the progress in molecular biology, it quickly became possible to transfer the gene responsible for producing the poisonous crystals into a plant. Corn became the first genetically modified plant to be officially registered with the Environmental Protection Agency (EPA) in 1995. Currently, genetically modified (GM) crops such as potatoes and cotton are cultivated worldwide.

In 1999*,* first time found that Cry proteins from three nonhemolytic & non-insecticidal Bt strains, named parasporin, exhibited high cytocidal activity in the leukemia T cells and many other human cancer cells (Mizuki *et al.*, 1999). Since then, 19 parasporins have been discovered. Recent advances in anticancer Cry proteins of Bt are shown in Figure 2.2.

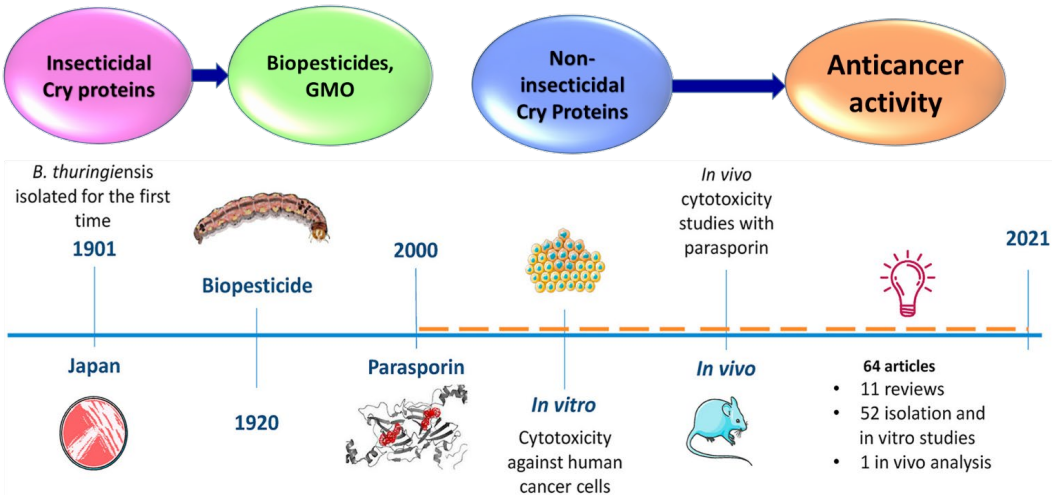



**Figure 2.2:** Recent advances in anticancer Cry proteins of Bt (Santos *et al.*, 2022).

## 2.3 Beneficial applications of *B. thuringiensis*

The biocontrol properties of *B. thuringiensis* (Bt) have made it one of the most well-known microorganisms in the agricultural industry (Jouzani *et al.,* 2017; Palma *et al.,* 2014). The use of *B. thuringiensis* as a biopesticide for the control of insect infestations has been extensively employed. Bt has replaced several chemical pesticides and insecticides because of their detrimental impact on the environment (Jallouli *et al.,* 2020; Kumar *et al.,* 2021). This is because Bt has been used against many invertebrate pests for decades for its diverse insecticidal activity and high specificity (Nicolopoulou-Stamati *et al.,* 2016). Researchers also looked at the antifungal properties of *Bt* (Reyes-ramírez *et al.,* 2004). Both *B. thuringiensis* and *B. thuringiensis* var. *israelensis* were utilized to study inhibitory effects. Bacteriocins from *B. thuringiensis* and *B. thuringiensis* subsp. *thuringiensis* was tested for stability and antibacterial activity by (Jung *et al.,* 2007). The non-insecticidal *B. thuringiensis* strains that produce parasporal inclusion were found to be more common than their insecticidal counterparts (Ichimatsu *et al.,* 2000; Meadows *et al.,* 1992; Mizuki *et al.,* 1999; Ohba, 1996; Ohba & Aizawa, 1988). Purified inclusion of three *B. thuringiensis* strains, 84-HS-1-11, 90-F-45-14, and 89-T-26-17, was shown to have a potent cytocidal effect against human leukemia T cells and other human cancer cell lines, as reported by Mizuki *et al.,* (1999). The beneficial applications of *B. thuringiensis* are shown in Figure 2.3.

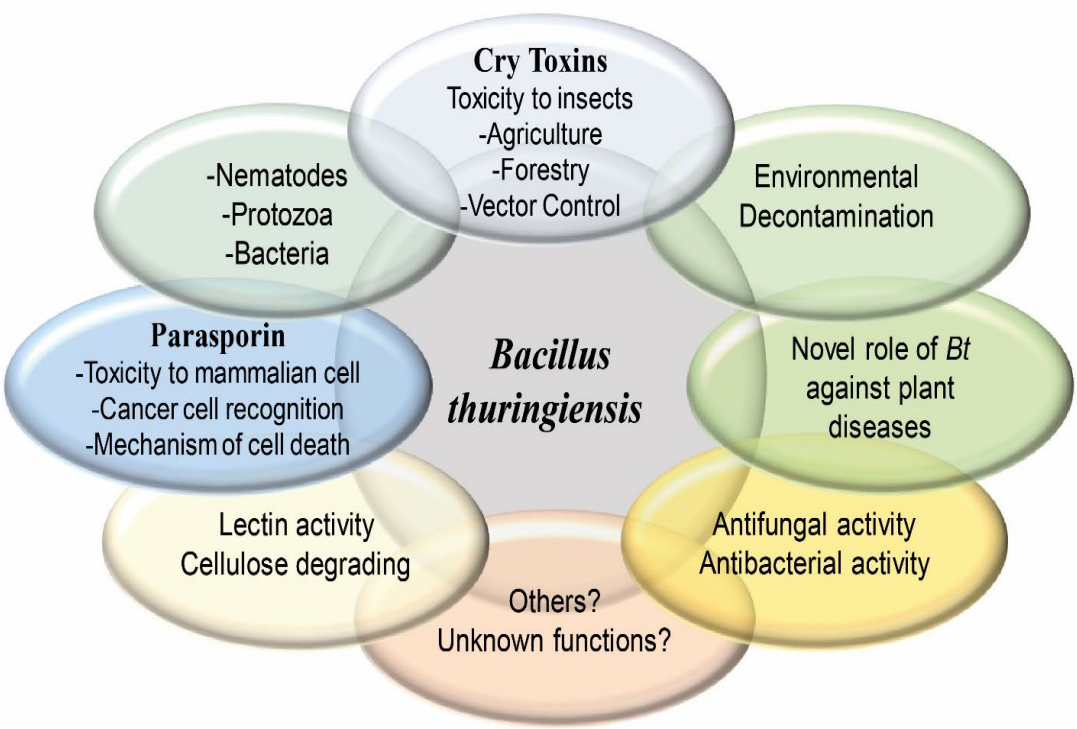



**Figure 2.3:** Beneficial applications of *B. thuringiensis* adopted and modified from Swamy & Asokan (2013).

## 2.4 Ecology and diversity of *B. thuringiensis*

*B. thuringiensis* is a member of a small group of naturally occurring bacteria that may also be introduced to an environment to control insects. Several studies have determined that *B. thuringiensis* is a naturally occurring, soil-dwelling bacteria that is found occasionally in almost every country on Earth. Researchers also isolated *B. thuringiensis* strains from various habitats, such as water, phylloplane, marine sediments, dead insects, grains, dust, etc. (DeLucca *et al.,* 1981; Kaelin *et al.,* 1994; Meadows *et al.,* 1992; Smith & Couche, 1991), grass foliage (Wang *et al.*, 1987), the intestine of mammals (Swiecicka *et al.,* 2002), grain storage, mushroom compost (Patel *et al.,* 2011). It had been isolated from several soil and plant types (Arrieta *et al.,* 2004; Bernhard *et al.,* 1997; Swamy & Asokan, 2013; Travers *et al.,* 1987) and mostly was found in phylloplane and soil habitats in vegetative form (Bizzarri & Bishop, 2008; Hendriksen & Hansen, 2002). The spores of the bacterium *B. thuringiensis* were found to easily endure in soil, and plant development happens when nutrients were present. Nonetheless, it was evident that *B. thuringiensis* also isolated from soil that had not been contaminated with *B. thuringiensis*-derived pesticides and in which there were

few sensitive insects. In addition, several of these isolates were harmless to all studied insect species. In 1986, (Ohba & Aizawa, 1986) obtained 189 strains of *B. thuringiensis* by isolating them from 136 soil samples collected from non-agricultural areas in Japan. The classification was partly determined by the existence of parasporal bodies, rather than insect pathogenicity. The vast majority of these *B. thuringiensis* isolates do not cause damage to various insect species (Ohba *et al.,* 1988). In 1981, De Lucca *et al.* conducted a similar study on the presence *of B. thuringiensis* in the United States. *B. thuringiensis* was found in a small proportion, ranging from 0.5% to 0.005%, of all *Bacillus* species that were obtained from soil samples. In 1987, Travers (Travers *et al.,* 1987) analyzed soil samples taken at random from Montgomery County, Maryland, USA, to determine the presence of *B. thuringiensis*. It was detected in almost all of the soil samples that were examined. Hence, it is apparent that *B. thuringiensis* is plentiful in the natural environment.

## 2.5 Classification of *B. thuringiensis*

*B. thuringiensis* (Bt), along with other *Bacillus* species, had been categorized based on its cellular, cultural, biochemical, and genetic properties (Carlson & Kolstø, 1993; Claus D, 1986; Hansen *et al.*, 1998; Lynch & Baumann, 1985; Slepecky & Hemphill, 2006). A new categorization system was developed and announced to classify different types of crystalliferous bacteria by looking at their morphological and biochemical characteristics (Heimpel, 1967; Heimpel & Aangus, 1958). However, recent molecular research suggested that there were several variations among serotypes and that certain biochemical characteristics didn't consistently align with a particular serotype (Hansen *et al.*, 1998; Helgason *et al.*, 1998). In 1999, Lecadet (Lecadet *et al.*, 1999) demonstrated that non-insecticidal serovars of the Bt subspecies display an escalating occurrence of abnormal crystal morphology, which includes the presence of novel crystal constituents. This finding was derived from the examination of δ-endotoxin profiles and the corresponding larvicidal action.

In the 1960s, for the first time the use of serological investigation of the flagella (H) antigens to categorize Bt subspecies were performed (Barjac & Bonnefoi, 1962). Furthermore, morphological and biochemical criteria have been used to augment this categorization, in addition to serotype classification (Barjac, 1981). Until 1977, a total

of 13 Bt subspecies had been recognized, and throughout that time, all subspecies were only toxic to lepidopteran larvae. The discovery of other subspecies that were toxic to Diptera (Goldberg & Margalit, 1977), Coleoptera (Krieg *et al.,* 1983), and perhaps Nematoda (Guo *et al.,* 2022) had broadened the range of organisms affected and substantially increased the number of subspecies. The hag gene of Bt encoded the genetic instructions for synthesizing flagellin, a protein that initiates the immunological response in H serotyping (Soufiane & Côté, 2009). The relationship between certain flagellin amino acid sequences and particular Bt H serotypes has been confirmed. The IEBC collection contained about 3500 isolates of *B. thuringiensis*. Among these isolates, researchers had found and defined at least 69 H serotypes, 13 sub-antigenic groups, and 82 serological variants (serovars) of Bt (Lecadet *et al.,* 1999). The H-serotyping technique was a simple and efficient method for classifying strains of *B. thuringiensis* species, which were identified by the presence of one or more parasporal inclusions. However, there were at least two situations that provided difficulties: (i) strains that didn't not contain a parasporal inclusion and were therefore categorized as *B. cereus*, and (ii) strains that demonstrate autoagglutination. The collection of *B. thuringiensis* strains has been continuously expanding, leading to a proportionate increase in the number of isolates falling into these two categories (Lecadet *et al.,* 1999). Due to its substantial economic importance, it is crucial to develop alternative approaches for classifying and grouping Bt strains and isolates.

## 2.6 Toxins of *B. thuringiensis*

Rowe *et al.* (1987) and the WHO in 1999 discovered a total of nine separate toxins in Bt strains. The bacteria produced various toxins, including: (i) α-exotoxin (phospholipase C), (ii) β-exotoxin (thermostable exotoxin), (iii) γ-exotoxin (toxic to sawflies), (iv) δ-endotoxin (parasporal crystal protein), (v) louse factor exotoxin (active only against lice), (vi) mouse factor exotoxin (toxic to mice and lepidopteran insects), (vii) water-soluble toxin, (viii) Vip3A (Bt vegetative insecticidal protein), and (ix) enterotoxin (produced by vegetative cells). The δ-endotoxin present in Bt strains exhibited toxicity against a wide variety of insect pests, nematodes, tumor cells, and cancer cells (Ekino *et al.,* 2014; Jouzani *et al.,* 2008; Kitada *et al.,* 2006; Mizuki *et al.,* 1999; Okumura *et al.,* 2008; Prasadi & Shethna, 1976; Ruan *et al.,* 2015).

## 2.7 Sporulation and parasporal inclusion formation of *B. thuringiensis*

*Bacillus thuringiensis,* when cultivated on a standard liquid medium, takes the form of straight rods of around 1.0-1.2 by 3-5 µm in size. Sporangia that were not enlarged produced ellipsoidal spores at their terminal or subterminal positions (Stahly *et al.,* 2006). *B. thuringiensis* has two distinct stages in its life cycle: the division of vegetative cells and the formation of spores. Midway down the plasma membrane, the vegetative cell forms a division septum, which eventually splits into two identical daughter cells (Bechtel & Bulla, 1982). Sporulation was distinguished by an unequal division of cells and encompasses seven different phases. The steps involved in this process are as follows (Figure 2.4): stage I, which involves the construction of an axial filament; stage II, which entails the creation of a forespore septum; stage III, which includes the process of engulfment, the first appearance of parasporal crystals, and the development of a forespore. Following that, stages IV to VI include the development of the exosporium, first cell wall, cortex, and spore coats, along with the alteration of the spore nucleoid. Finally, stage VII, the original mother cell undergoes lysis, resulting in the release of the endospore and parasporal inclusion.

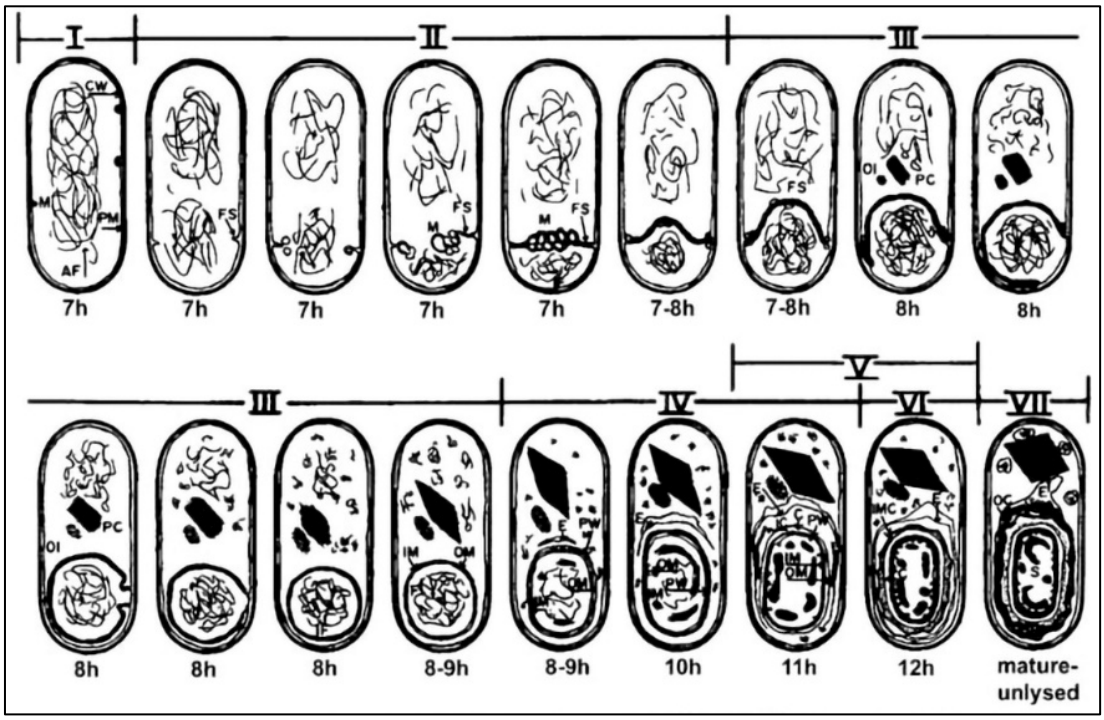



**Figure 2.4:** Diagram illustrating the sequential process of sporulation and parasporal inclusion formation in *B. thuringiensis* (Bechtel & Bulla, 1976).

## 2.8 Morphological diversity of parasporal Crystalline inclusions

A notable characteristic of the Bt bacterium is its ability to produce parasporal crystalline inclusions during the stationary phase of sporulation (Höfte & Whiteley, 1989; Schnepf *et al.*, 1998). The existence of parasporal inclusions in Bt was first noted in 1915 (Berliner, 1915), but the precise protein content of these inclusions was not determined until the 1950s (Angus, 1954). Hannay (1953) noted the existence of a crystalline fine structure in most of the parasporal inclusions. In 1967, Heimpel used the name "parasporal crystalline protein" to describe the toxic substance present in the cell of Bt. Several subspecies of Bt possess the capability to generate diverse types of inclusions, which might be comprised of unique protein compositions. The form of the parasporal crystals may exhibit variation based on the amino acid composition of the protein. These crystals may assume many geometric configurations, including bipyramidal, cuboidal, circular, flat rhomboid, spherical, square, or irregular shapes. Occasionally, the crystals might be a composite of two or more protein varieties (Chilcott & Wigley, 1993, 1994). Bipyramidal crystals had shown efficacy against lepidopteran species (Attathom *et al.,* 1995). Cuboidal crystals exhibited efficacy against both lepidopteran and dipteran larvae, or exclusively against lepidopteran larvae (Yamamoto & McLaughlin, 1981). Spherical and irregular crystals predominantly exhibited mosquitocidal capabilities and often exhibited action against certain coleopteran species (Krieg *et al.*, 1983). Irregular crystals showed little or no detectable toxicity (Zelazny *et al.,* 1994). Furthermore, some Bt produced round or circular inclusions that possess surface layer homology domain (Guo *et al.,* 2008; Sun *et al.,* 2001; Zhu & Yu, 2008). A correlation had been found between the morphology of crystals, the presence of insecticidal crystalline proteins (ICPs), and their efficacy against certain insects (Bulla *et al.,* 1977; Höfte & Whiteley, 1989; Lynch & Baumann, 1985). Figure 2.5 exhibits scanning and transmission electron microscopic images showing several crystal morphologies of Bt strains.

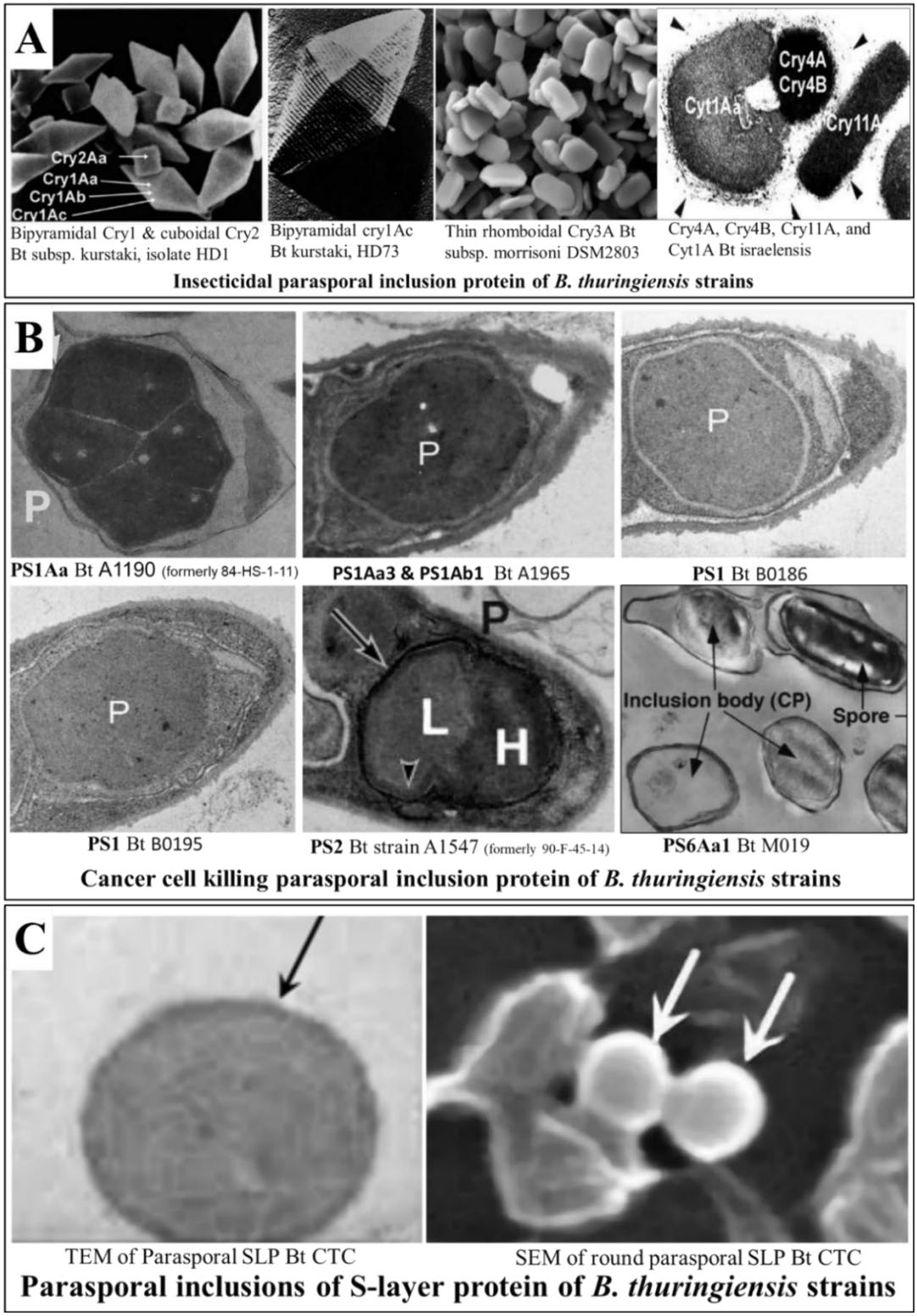



**Figure 2.5:** Scanning and transmission electron microscopic image displaying different crystals shape of Bt strains. (A) insecticidal inclusions (Federici, 2005; Sawaya *et al.*, 2014); (B) anticancer inclusions (Mizuki *et al.,* 2000; Kim *et al.,* 2000; Uemori *et al.,* 2007; Nagamatsu *et al.,* 2010); (C) S-layer inclusions (Zhu & Yu, 2008).

## 2.9 Properties of parasporal Crystalline inclusions in *B. thuringiensis*

The parasporal crystalline inclusions of *B. thuringiensis* consist mostly of protein molecules. The protein molecules mostly consist of three distinct types: (i) Cry; (ii) Cyt; and (iii) surface layer (S-layer) protein. Of them, Cry's and Cyt's were classified as delta endotoxin family proteins, whereas S-layer proteins were identified as two-dimensional crystalline arrays of cell surface proteins or glycoprotein subunits. S-layers were found to be connected with the cell membrane (Zhu *et al.,* 2017). Two different types of parasporal S-layer proteins were identified, one of them was a surface array protein (Sap) and other was an extractable antigen 1 (EA1) protein.

## 2.10 Solubilization and activation of Cry toxin

The crystal toxin proteins were first synthesized in an inactive state they were not soluble in regular water or inorganic solvents. Nevertheless, they have the potential to undergo dissolution in alkaline solutions that possess a pH ranging from 8.5 to 12. . The Cry2 proteins exhibited solubility at a pH of around 12, while the Cry1 proteins exhibited solubility at a pH of 9.5. Similarly, the toxins Cry4A, Cry5B, and Cyt exhibited solubility at a pH of 9.5, but the toxin Cry4D required a pH of 12. The Cry3A toxin had a solubility that decreased at pH levels below 4 and above 9.5 (Koller & Hollingworth, 1992). Typically, Cry toxin proteins were found to be dissolved in the alkaline conditions seen in the midgut of insect larvae (Tojo & Aizawa, 1983). The crystals became solubilized under neutral pH conditions in the presence of detergents and denaturing agents such as urea, β-mercaptoethanol, dithiothreitol (DTT), and SDS. Solubilized Cry proteins became activated by digestion by proteases found in the midgut of insects (Choma *et al.*, 1990; Tojo & Aizawa, 1983) or through artificial treatment with proteinase K or trypsin (Ishii & Ohba, 1993). This process of activating protoxins were widely recognized as a crucial and fundamental step in exerting toxicity (Figure 2.6).

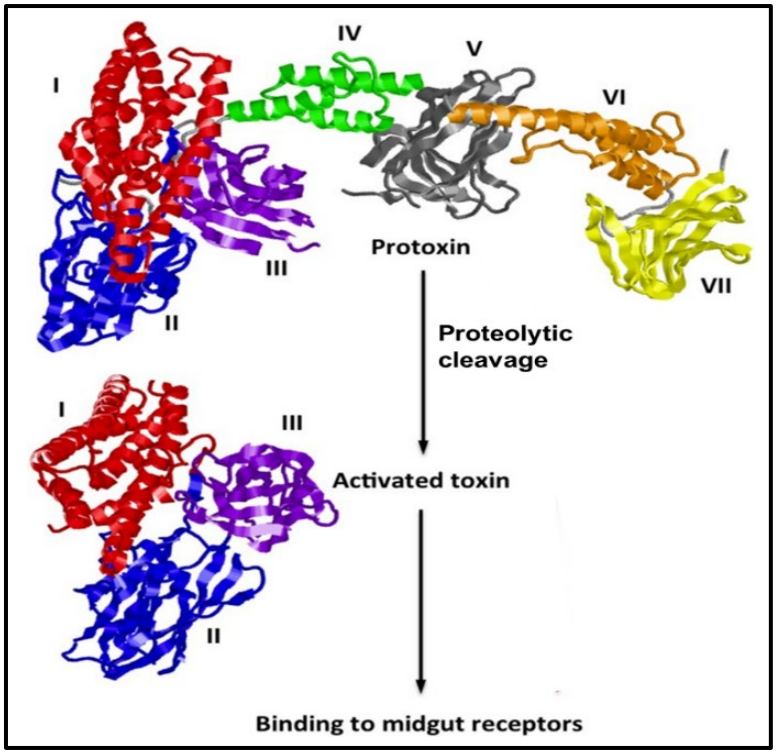



**Figure 2.6:** Bt protein activation. In the traditional paradigm, the inactive Cry1Ac protoxin (consisting of domains I-VII; PDB 4W8J) has to be transformed into the activated toxin (consisting of domains I-III; PDB 4ARY) before it can attach to insect midgut receptors and cause harm (Tabashnik *et al*., 2015).

## 2.11 The diversity of Cry toxins in *B. thuringiensis* strains

The word Cry is derived from the presence of proteins inside a parasporal crystal. During sporulation, they synthesized as as protoxins and were then broken down at the N- and/or C- terminal to create a poisonous core (Schnepf *et al.*, 1998). Crystals in some Bt strains may make up to 25% of the dry weight of the sporulated cell (Agaisse & Lereclus, 1995). Most of the Cry toxins were located in parasporal crystals(Aronson *et al.,* 1982). The genes that encoded Cry toxins were believed to be located on extensive plasmids, with many Bt strains possessing several cry genes (Carlton and Gonzalez, 1984). There was several scientific evidence that Cry protein showed toxicity against several organisms, including lepidopteran moths and butterflies, dipteran flies, coleopteran beetles, hemipteran cicadas, aphids, and bugs, hymenopteran bees, ants, and wasps, nematodes, snails, and cancer cells (Palma *et al.,* 2014) (Figure 2.7).

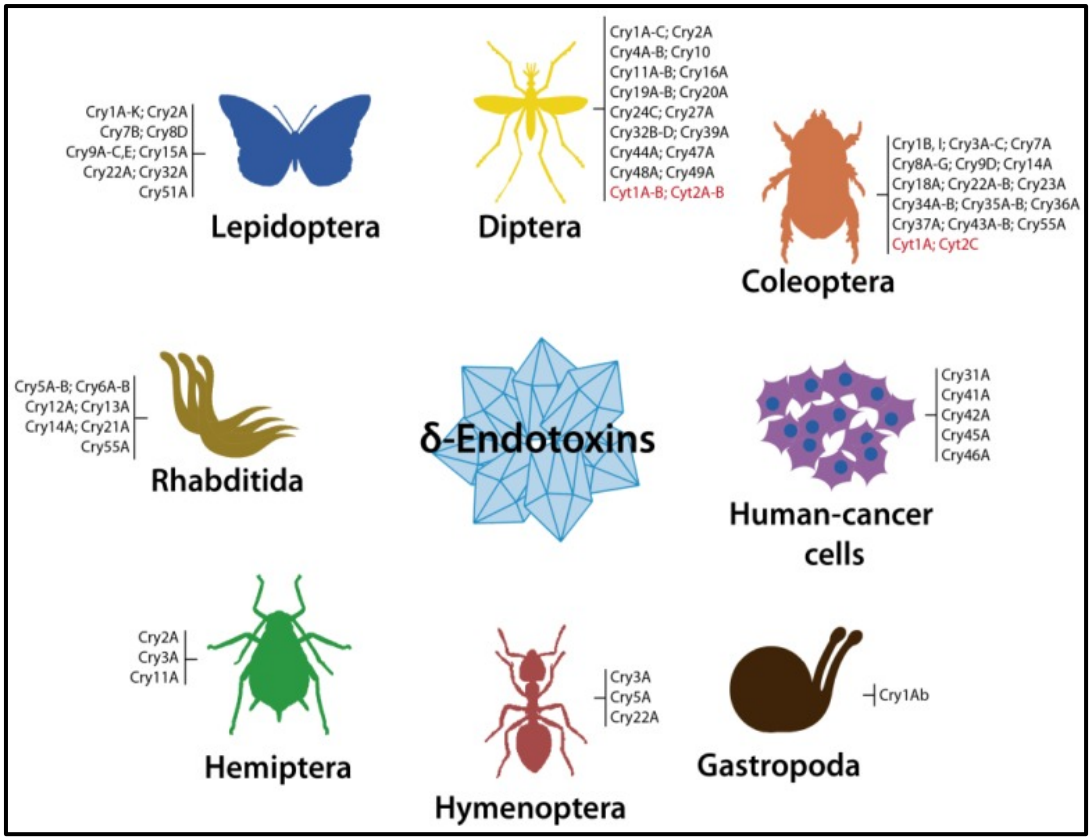



**Figure 2.7:** Different types of well-known hosts that Bt δ-endotoxins (Cry and Cyt) affects. This understanding was based on the selectivity of the three-domain toxins towards various insect orders (Frankenhuyzen, 2009; Frankenhuyzen, 2013). The diagram is taken from Palma *et al.,* (2014).

Currently, a documentation in 2014 identified around 75 fundamental subgroups of Cry toxins (Adang *et al.,* 2014). The subgroups were differentiated by unique main ranks in the nomenclature, which spanned from Cry1 to Cry75. Figure 2.8 is an illustrative presentation that depicts the variation of the lengths of these toxins. Their lengths ranged from 1344 to 369 amino acids, namely Cry34 and Cry43. In 1989, Höfte and Whiteley identified five conserved sequence blocks that are universally present in all the Cry toxins. Some toxins only include a subset of these components, and not all toxins contain them.

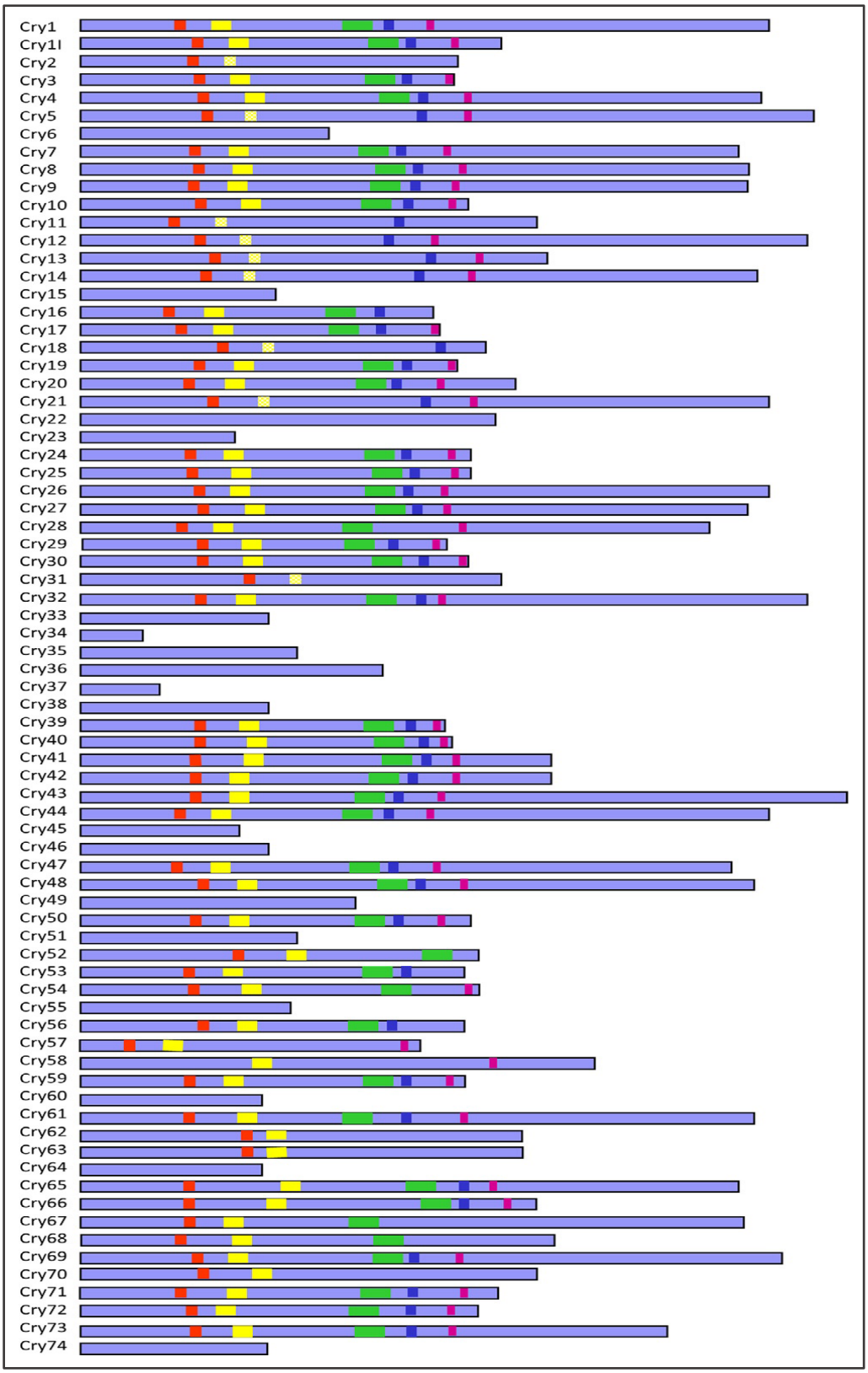



**Figure 2.8:** Diagrammatic representations of the range of Bt Cry toxins.

The length of each toxin is clearly shown, and the five conserved blocks are presented as colored inserts (Adang *et al.*, 2014).

## 2.12 The 3D crystallographic structure of the three-domain cry toxin

The three-domain Cry proteins are the largest and most varied group of insecticidal proteins found in Bt parasporal crystalline inclusions. The three-domain Cry toxins have been extensively used as biological insecticides for a significant duration. The key attribute of cry proteins is in their diversity, irrespective of the precise arrangement of their amino acid sequences, since they all possess a consistent three-domain 3D structure. The crystal structures of Cry1Aa, Cry2Aa, Cry3A, Cry3Bb, Cry4Aa, Cry4Ba, and Cry8Ea1 had been determined using X-ray crystallography (Boonserm *et al.,* 2005; Boonserm *et al.,* 2006; Galitsky *et al.,* 2001; Grochulski *et al.,* 1995; Guo *et al.,* 2009; Li *et al.,* 1991; Morse *et al.,* 2001). An illustration of the three-domain structure is shown in Figure 2.9.

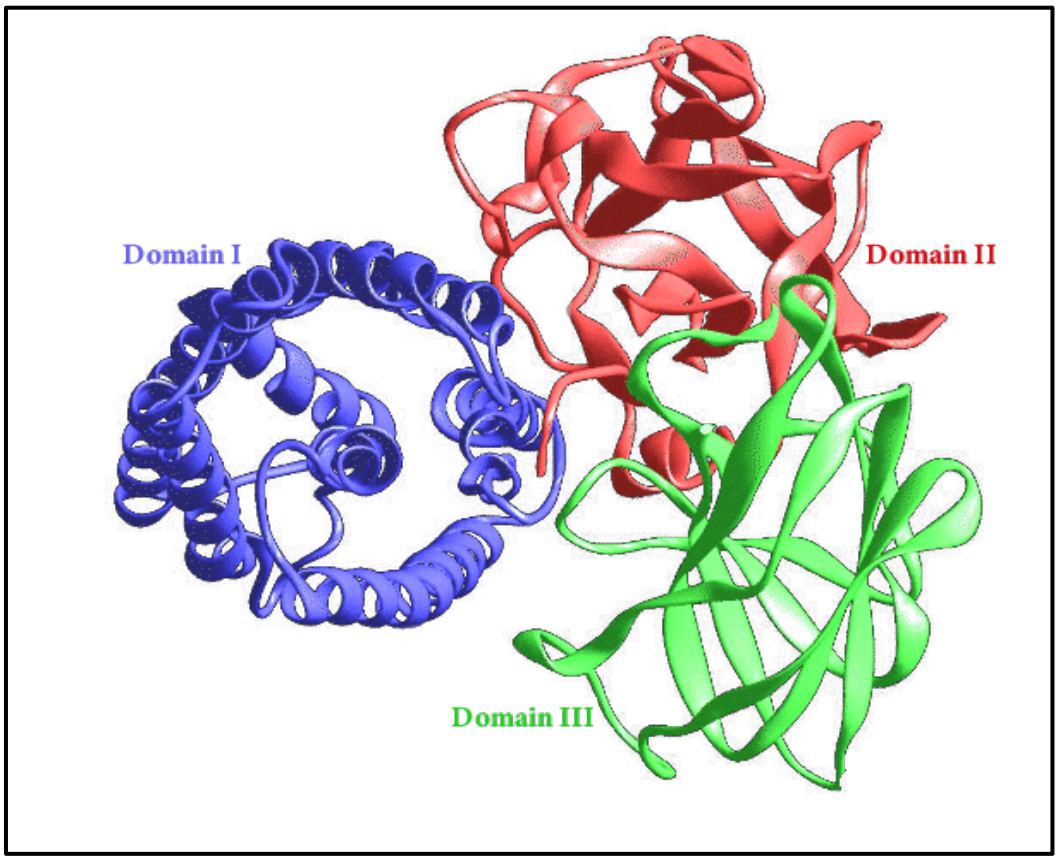



**Figure 2.9:** The 3-D structure of Bt Cry3A protein. RCSB-PDB id 4QX1 (Sawaya *et al.*, 2014). The three-domain Cry proteins are often represented by Roman numerals: (I) is the perforating domain; (II) is the central domain that is engaged in toxin-receptor interactions; and (III) is the galactose-binding domain that is involved in receptor binding and pore creation.

**Domain-I,** or the perforating domain, situated at the N-terminus, consists of 7-8 α-helices, including a central hydrophobic α-helix. It has a role in the process of membrane insertion and pore formation (Li *et al.,* 1991). The provided theoretical framework proposes that α-helices 4 and 5 penetrate the membrane by adopting a hydrophobic hairpin conformation, whereas the other α-helices are oriented parallel to the cell membrane. (Li *et al.,* 1991).

**Domain-II,** or the receptor binding domain, consists of three antiparallel β-sheets and has a significant function in selectivity and receptor binding (Schnepf *et al.*, 1998). The base of domain II contains three flexible loops, where the second and third loops play a key role in recognizing and attaching to receptors (Adang *et al.*, 2014).

**Domain-III,** or the galactose-binding domain, consists of two antiparallel β-sheet layers arranged in a sandwich-like structure. This domain is responsible for receptor binding and may play a role in pore creation as well (Li *et al.,* 1991). The structure has a resemblance to a galactose binding domain and encompasses conserved blocks 3-5 (Adang *et al.*, 2014).

## 2.13 Cyt Toxins

The cytolytic toxins were first identified in Bt subspecies israelensis (Goldberg & Margalit, 1977) and have recently been examined in detail (Ben-Dov, 2014; Soberón *et al.,* 2013). These proteins resemble to be exclusive to certain Bt subspecies (Tyrell *et al.*, 1981). The 28 kDa protein derived from Bt subsp. israelensis was first characterized by its ability to cause cell death in both insect and mammalian cells grown in culture, as well as its toxic effects when administered to mice by injection. The 28-kDa protein enhanced the effectiveness of other Bt subsp. israelensis crystal proteins, but it exhibited lower efficacy against mosquitoes compared to native Bt subsp. israelensis crystals (Chilcott & Ellar, 1988; Davidson & Yamamoto, 1984; Wu & Chang, 1985). These genes belong to three cyt toxin gene families (*cyt1, cyt2*, and *cyt3*) and comprise a total of 11 holotype toxins according to the current nomenclature.

The Cyt1Aa (Cohen *et al.*, 2011) and Cyt2Ba (Cohen *et al.*, 2008) proteins had been seen to have three-dimensional structures that consist of a single domain and three layers of alpha-beta proteins (Figure 2.10). According to mutagenic experiments, β-

sheet residues were essential for toxicity, but helical domain residue mutations did not affect toxicity, indicating a crucial function for the β-sheet core (Butko, 2003).

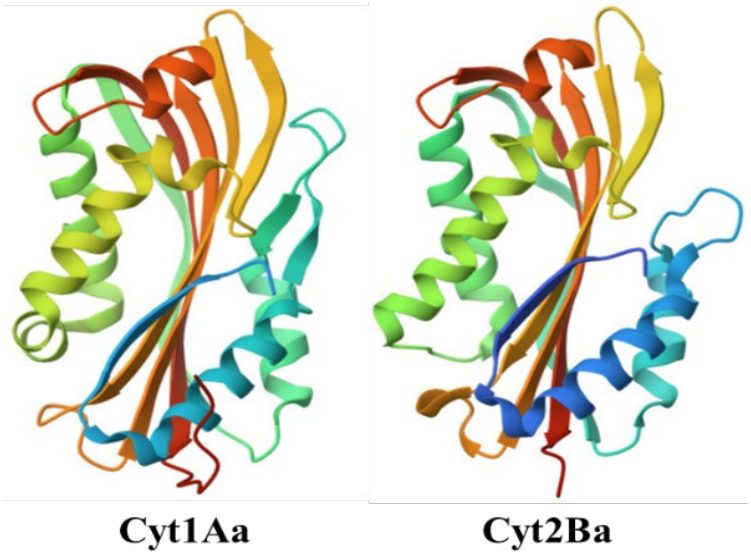



**Figure 2.10:** The 3-D structure of the toxin Cyt1Aa and the activated Cyt2Ba monomer. RCSB PDB accessions: 3RON (Cohen *et al.*, 2011) and 2RCI (Cohen *et al.*, 2008), respectively.

Digestive proteases at the identical locations in the N- and C-termini break down ingested Cyt1A and Cyt2A protoxins to produce stable toxin cores that are 25 and 23 kDa, respectively (Konit & Ellar, 1994). The midgut brush border membrane of Diptera contains a lot of membrane lipids with unsaturated acyl chains, for which activated cyst toxins have a strong affinity (Gill *et al.,* 1987; Li *et al.,* 1996; Thomas & Ellar, 1983). It was found that Cyt2Ba has a possible pocket for phospholipids to attach, which is similar to the pocket in the Erwinia virulence factor (Evf) (Rigden, 2009). Butko (2003) suggested two distinct mechanisms for Cyt1A-induced cytolysis: hole creation and detergent-like membrane rupture. The findings from experiments using planar bilayers (Knowles *et al.*, 1989) and erythrocytes (Promdonky & Ellar, 2003), as well as the significance of pre-pore oligomerization in terms of toxicity (Soberón *et al.*, 2013), supported the development of pores by Cyt toxins.

## 2.14 Parasporin (PS), the anticancer Cry toxin of *B. thuringiensis*

In addition to the insecticidal Cry toxins, researchers also discovered non-insecticidal, and nonhemolytic Cry toxins from Bt strains that specifically cytotoxic to human cancer cell lines but noncytotoxic to normal cell line (Kim *et al.*, 2000; Lee *et al.*, 2000; Mizuki *et al.*, 2000; Yamashita *et al.*, 2000). Mizuki *et al., (*2000) coined the name parasporin for the cancer cell killing parasporal inclusion forming protein of Bt. In addition to parasporins, cytotoxic effects of parasporal inclusion forming protein were also obtained from two mosquitocidal reference strains, *israelensis* and *kyushensis*, on mammalian cells, fish, and insect cells (Al-yahyaee & Ellar, 1995; Grisolia *et al.,* 2009; Kim *et al.,* 2000; Knowles *et al.,* 1989; Thomas & Ellar, 1983). Additionally, the cytocidal action also had been ascribed to the Cyt toxins (Thomas and Ellar, 1983; Mizuki *et al.*, 1999).

Parasporins were identified as genealogical heterogenous (Mizuki *et al.,* 2000) hence a committee was established in 2006 for proper classification and nomenclature of parasporin based on the similarity of amino acid sequences. The Cry gene nomenclature was expanded to include all Bt parasporins because of their high sequence resemblance to known insecticidal toxins. The parasporin had been classified into six primary classes, from parasporin-1 to 6 or PS1-PS6 (Figure 2.11). Parasporins are very rare in Bt and there are 19 different parasporins discovered to date (Table 2.1).

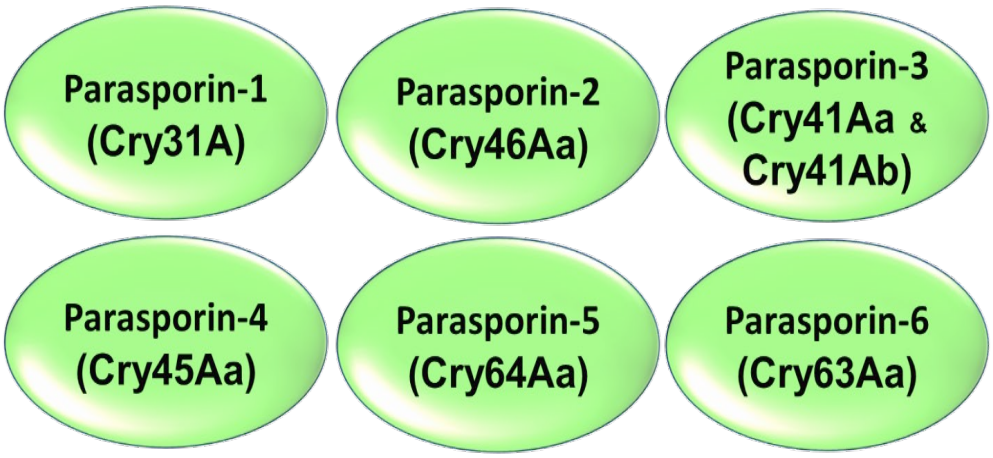



**Figure 2.11:** Diagrammatic presentation of six main classes of Cry proteins belonging parasporins

**Table 2.14:** A list of all 19 discovered parasporins with brief descriptions

| PS | Bt Strain | Cry Gene | Protoxin (kDa) | Active Toxin (kDa) | Protease Activation | Main Cellular Target | EC50 [µg/mL] | Mechanism Action | Country |
|---|---|---|---|---|---|---|---|---|---|
| PS1Aa1 | A1190 | Cry31Aa1 | 81 | 15, 56 | Trypsin | HeLa | 0.12 | Apoptosis | Japan |
| PS1Aa2 | M15 | Cry31Aa2 | 83 | 55, 70 | Trypsin | Jurkat | 0.02 | ND | Canada |
| | | | | | | HepG2 | 0.02 | | |
| PS1Aa3 | B195 | Cry31Aa3 | 81 | 56 | Trypsin | HeLa | 14.7 | ND | Japan |
| PS1Aa4 | Bt 79-25 | Cry31Aa4 | 81 | NP | Proteinase K | ND | ND | ND | Vietnam |
| PS1Aa5 | Bt 92-10 | Cry31Aa5 | 81 | NP | Proteinase K | ND | ND | ND | Vietnam |
| PS1Aa6 | M019, | Cry31Aa6 | 70 | 15, 55 | Trypsin | HepG2 | 0.52 | ND | Japan, |
| | CP78A | | | | | | | | Caribbean |
| PS1Ab1 | B195 | Cry31Ab1 | 82 | 56 | Trypsin | HeLa | 14.7 | ND | Japan |
| PS1Ab2 | Bt 31-5 | Cry31Ab2 | 82 | NP | Proteinase K | ND | ND | ND | Vietnam |
| PS1Ac1 | Bt 87-29 | Cry31Ac1 | 87 | NP | Proteinase K | ND | ND | ND | Vietnam |
| PS1Ac2 | B0462 | Cry31Ac2 | 81 | 15, 60 | Proteinase K | HeLa | 2 | Apoptosis | Japan |
| PS1Ad1 | CP78B, M019 | Cry31Ad1 | 73 | 14, 59 | Trypsin | HepG2 | 0.52 | ND | Caribbean, Canada, Japan |
| PS2Aa1 | A1547 | Cry46Aa1 | 37 | 30 | Proteinase K | HepG2 | 0.023 | Pore-forming | Japan, USA |
| PS2Aa2 | A1470 | Cry31Aa2 | 30 | 28 | Proteinase K | MOLT-4 | 0.041 | ND | Japan |
| PS2Ab1 | TK-E6 | Cry31Ab1 | 33 | 29 | Proteinase K | Jurkat | 0.545ng/ml | ND | Japan |
| PS3Aa1 | A1462 | Cry41Aa1 | 88 | 64 | Proteinase K | HL60 | 1.32 | ND | Japan |
| PS3Ab1 | A1462 | Cry41Ab1 | 88 | 64 | Proteinase K | HL60 | 1.25 | ND | Japan |
| PS4Aa1 | A1470 | Cry45Aa1 | 31 | 28 | Proteinase K | CaCo2 | 0.124 | Pore-forming | Japan |
| PS5Aa1 | A1100 | Cry64Aa1 | 33 | 30 | Proteinase K | TCS | 0.046 | ND | Japan |
| PS6Aa1 | M019, | Cry63Aa1 | 85 | 14, 59 | Trypsin | HepG2 | 2.3 | ND | Japan, |
| | CP84 | | | | | | | | Caribbean |

### 2.14.1 Parasporin-1

Parasporin-1 (PS1) was identified as an anticancer toxin (Cry31A) derived from *B. thuringiensis* that consists of three domains which didn't not have insecticidal and hemolytic properties. PS1 had a strong preference for causing cell death in human cancer cell lines. PS1 was found to produce polygonal parasporal inclusions (Figure 2.5 B). The cytocidal action of PS1 was seen only when it was broken down by proteases into smaller molecules ranging from 40 to 60 kDa. PS1 was activated by trypsin and proteinase K, but not by chymotrypsin, once activated, it had potent cytotoxic effects on human leukemic T cells (MOLT-4) and human uterine cervix cancer cells (HeLa). However, it did not demonstrate any cytotoxicity towards normal T cells (Mizuki *et al.,* 2000) (Figure 2.12). Researchers obtained PS1 from ten different strains of B. thuringiensis found in Japanese and Vietnamese soils, as well as from dead two-spotted spider mites parasitic on orchard apples in Canada (Jung *et al.,* 2007; Uemori *et al.,* 2007, 2008). PS1 was found to synthesize as an 81-kDa protein. The protein sequence of this toxin, which belongs to three domains, showed little similarity to the existing Cry and Cyt proteins. Researchers found that PS1 reduces the amount of cellular protein and DNA synthesis in HeLa cells(Katayama *et al.*, 2007). The first alteration seen in cells exposed to this toxin was a rapid rise in the intracellular free-Ca2+ concentration. This enhancement in the intracellular Ca2+ levels was noted 1–3 minutes after the administration of PS1 (Katayama *et al.*, 2007).

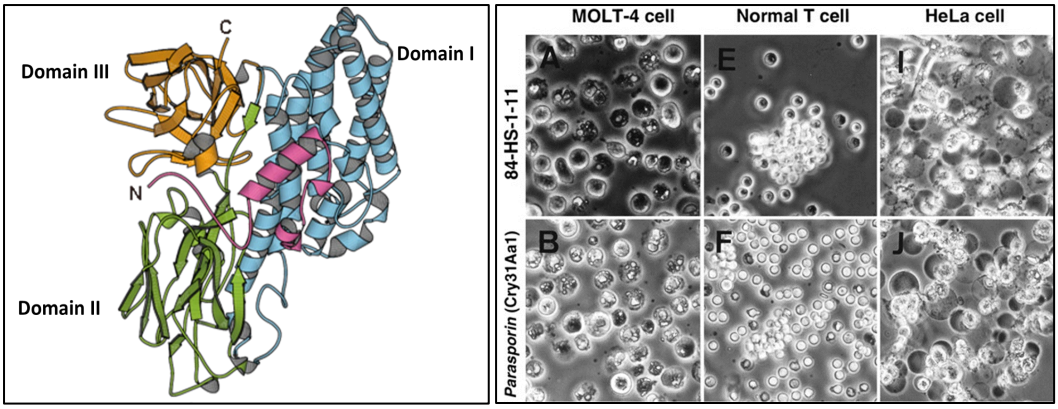



**Figure 2.12:** Three-domain PS1 Cry protein (left side) and cytotoxic effect of PS1 of *B. thuringiensis* strain 84-HS-1-11 (right side). Adopted from Mizuki *et al.* (2000)

### 2.14.2 Parasporin-2

Parasporin-2, also called Cry46Aa, is a toxin similar to hydralysin specifically attacks cancer cells in the liver and colon. PS2 was found to be produced in the form of parasporal inclusions (Figures 2.1 A and 2.5 B). PS2 didn't correlate with the majority of *B. thuringiensis* Cry or Cyt toxins. Proteinase-K's proteolytic cleavage of the 37 kDa PS2 produced a cytotoxic 30 kDa Cry protein fragment with potent activity against specific types of human cancer cells Enzymatic digestion eliminated 51 amino acid residues from the N-terminal end and 36 residues from the C-terminal end of the protein. Protease-activated PS2 caused notable changes in cellular morphology, such as cell blebbing, cytoskeletal abnormalities, and fragmentation of the mitochondria and endoplasmic reticulum. The cells exposed to intoxication showed a continued presence of mitochondrial and endoplasmic reticulum proteins. However, the permeability of the plasma membrane increased rapidly, leading to the release of the majority of cytoplasmic proteins. Researchers found that the PS2 exhibited preferential cytotoxicity against HepG2 and Jurkat cells, while exhibiting reduced toxicity towards HeLa and normal hepatocyte cells (Figure 2.13) (Ito *et al.,* 2004; Kitada *et al.,* 2006). The non-apoptotic cytotoxicity of PS2 leads to the swelling and fragmentation of susceptible cells, potentially impacting their ion permeability. The active PS2 crystal structure revealed an elongated shape that was characterized by the alignment of its long β-strands along its longitudinal axis. The PS2 was identified to have three domains. A small α-helix is positioned between two small β-sheets, which is likely responsible for the formation of the target binding module (Figure 2.13). Domains 2 and 3 are β sandwiches that might have a role in the process of oligomerization and the development of pores (Akiba *et al.*, 2009).

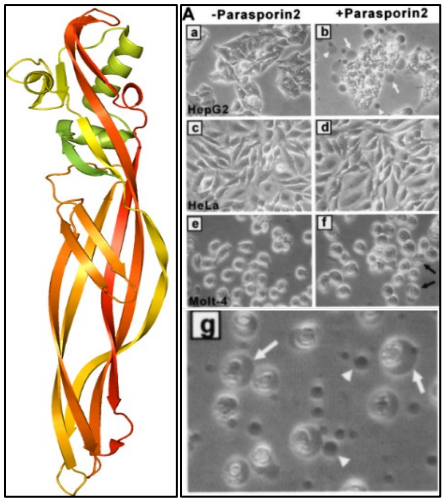



**Figure 2.13:** The 3D structure of activated PS2 and cytotoxic effects on cancer cell. In the left side the 3D structure of PS2 (PDB: 2ZTB), and in the right, morphological changes in HepG2, HeLa, and MOLT-4 cells due to PS2 (Kitada *et al.*, 2006).

### 2.14.3 Parasporin-3

Cry41Aa and Cry41Ab represented parasporin-3. An 88-kDa PS3 protein was isolated from *B. thuringiensis* strain A1462 from Tokyo soils (Yamashita *et al.,* 2005). The 64-kDa proteinase K-activated protein was preferentially cytotoxic to HepG2 (liver) and HL60 (myeloid leukemia cells) (Figure 2.14). The amino acid sequence of PS3 was comparable to the insecticidal Cry proteins of *B. thuringiensis*, but it selectively killed human cancer cell lines (Yamashita *et al.,* 2005). Little is known about it.

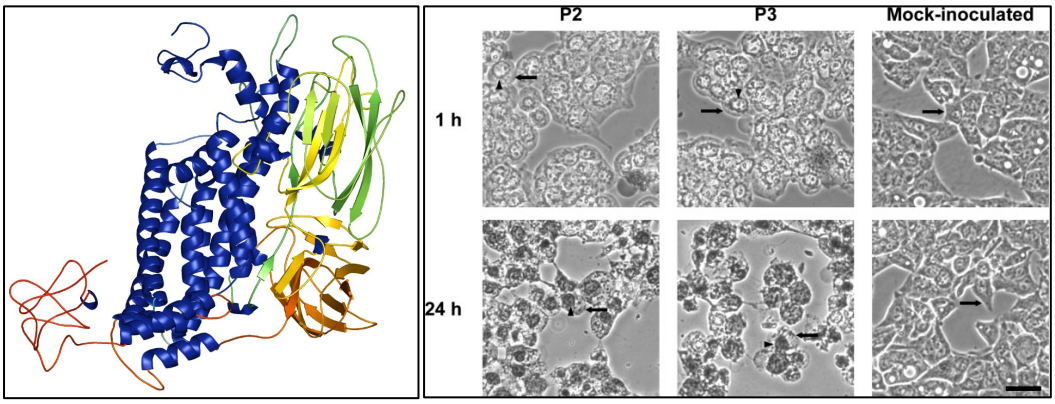



**Figure 2.14:** Modelled 3D structure of PS3 and cytotoxic effects on cancer cell. In the left the modelled 3D structure of PS3, and in the right, phase-contrast microscopy showed cytotoxic effects of PS3 toxins on HepG2 cells (Yamashita *et al.,* 2005)**.**

### 2.14.4 Parasporin-4

*B. thuringiensis* strain A1470 produced cytotoxic PS4 (Cry45Aa). The 31 kDa pro-parasporin-4 was changed into a 27 kDa active toxin when pepsin cut the C-terminus at Glu-252 (Okumura *et al.*, 2014). Researchers found that the activated PS4 toxin protein killed CACO-2, Sawano, MOLT-4, TCS, and HL60 cells (Figure 2.15) (Okumura *et al.*, 2013; Okumura *et al.*, 2011). Furthermore, the study revealed no cytotoxicity against human primary hepatocyte cells, UtSMC, MRC-5, or normal T cells (Okumura *et al.*, 2011). Researchers found that PS4 therapy inflates and shrinks nuclei, bursting CACO-2, Sawano, MOLT-4, TCS, and HL60 cells within 24 hours(Okumura *et al.*, 2011). Researchers identified PS4 as similar to the Epsilon/Mtx-like cytotoxic protein, primarily a -sheet protein (Figure 2.15). PS4 also had similarities with the pore-forming ε-toxin of *Clostridium perfringens.*

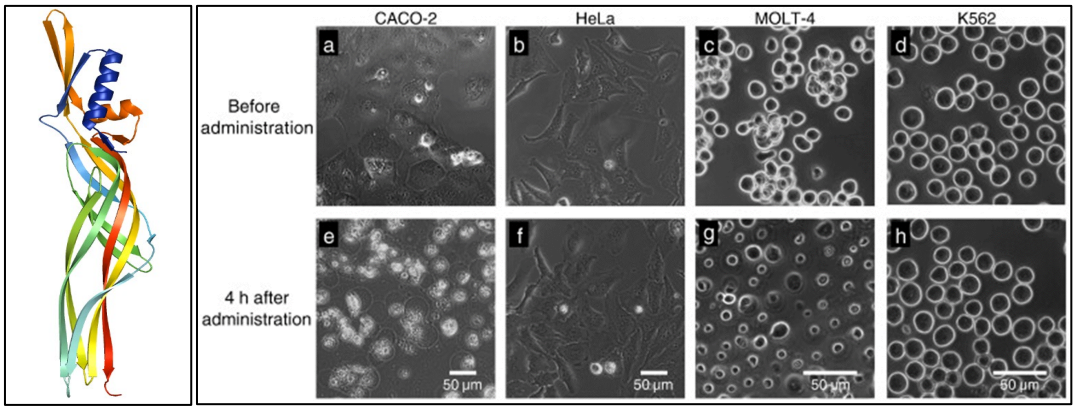



**Figure 2.15:** The 3D structure of activated PS4 and cytotoxic effects on cancer cell line. In the left side, the 3D structure of PS4 (PDB: 2D42), and in the right side, morphological changes in CACO-2, HeLa, MOLT-4 and K562 cell line due to toxicity of activated PS4 (Okumura *et al.*, 2011).

### 2.14.5 Parasporin-5

Bt strain A1100 (Ekino *et al.,* 2014) synthesized Parasporin-5, or Cry64Aa, as a 33.8-kDa inactive precursor protein. It was cytocidal against human leukemic T cells (MOLT-4) only when proteinase K at the C-terminal degraded it into 29.8-kDa molecules (Ekino *et al.*, 2014) (Figure 2.16). PS5 did not have a lot in common with other parasporins, and PSI-BLAST showed that it had some connections to *B. thuringiensis* Cry toxins and aerolysin-type β-pore-forming toxins (β-PFTs). Sequence analysis and modelled 3D structure suggest that it is a β-sheet-dominated protein. (Figure 2.16).

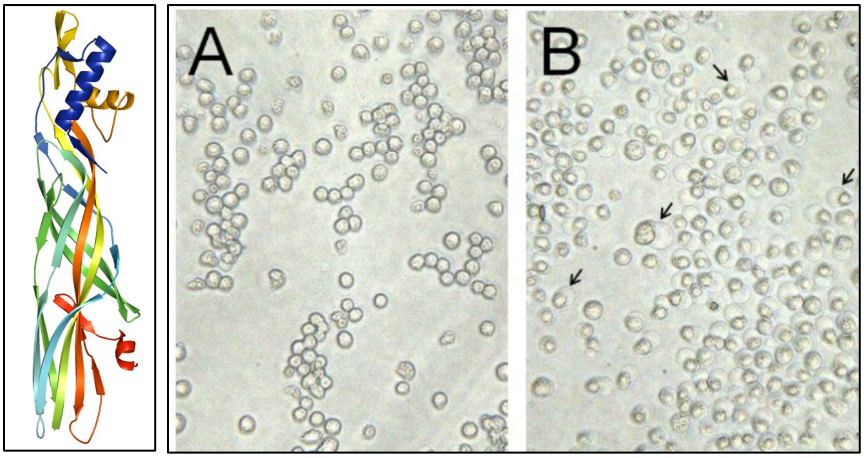



**Figure 2.16:** Modelled 3D structure of PS (left side). PS5 of *B. thuringiensis* A1100 showed cytotoxicity on MOLT-4 cells (right side). (A) Mock-inoculated cells; (B) cell incubation with solubilized protein at 37 °C for 1 h. The arrows represent cytotoxic protein-induced cell swelling (Ekino *et al.*, 2014).

### 2.14.6 Parasporin-6

Parasporin-6, or Cry63Aa, was extracted from *B. thuringiensis* strains M109 and CP84 (Nagamatsu *et al.,* 2010). The strains M109 and CP84 produced PS6 as unconventional oval-shaped parasporal inclusions (Figure 2.5 B), and sequence analysis revealed it to be a Cry toxin with three distinct domains (Figure 2.17). The trypsin digested PS6 toxin exhibited cytotoxic effects on HepG2 cells and HeLa cells (Figure 2.17) (Nagamatsu *et al.,* 2010).

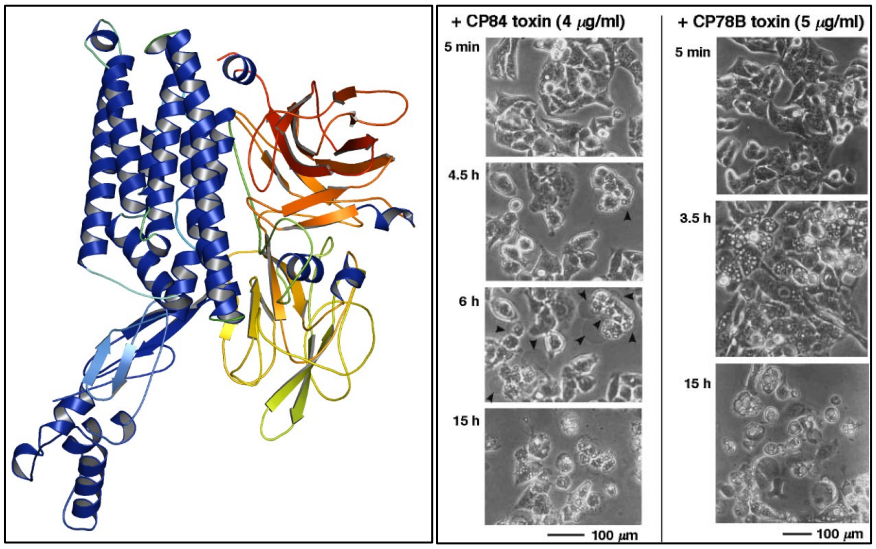



**Figure 2.17:** Modelled 3D structure of parasporin 6 (left side) and cytotoxicity of activated PS6 on HepG2 cells (right side). Cytotoxicity image adopted from Nagamatsu *et al.*, (2010).

### 2.14.7 Receptors of parasporins

**Beclin-1 and GAPDH**: In the case of parasporin-1, the identification of Beclin-1 as a PS-1 receptor in HeLa cells was reported (Patent No. EP2273266A1, 2011). Beclin-1 is a protein known for its tumor suppressor function. Furthermore, the researchers discovered that glyceraldehyde-3-phosphate dehydrogenase (GAPDH) acts as a binding protein on the plasma membrane of CEM-SS cells for the parasporin-like protein Bt18 (Krishnan *et al.,* 2010).

**Cholesterol and GPI-anchored proteins:** The receptors for parasporin 2 exhibited a little reliance on cholesterol, but they need GPI-anchored proteins. These proteins were mostly found in lipid rafts that were loaded with cholesterol and sphingolipids (Kitada *et al.,* 2009; Legler *et al.,* 2005; Lingwood *et al.,* 2010). The glycan part of the GPI anchor might make it easier for parasporin-2 to stick to the surface and then fit together with the membrane (Kitada *et al.*, 2009).

**Plasma membrane:** The plasma membrane was not selectively a target for parasporin-4. This caused a multi-subunit pore complex to form inside target cells. During this process, it exhibited a cholesterol-independent action that was different from parasporin-2 (Kitada *et al.*, 2009; Okumura *et al.*, 2011). Parasporin-4 may function as a distinct cholesterol-independent β-PFT. However, more investigation is required to determine the particular receptor and its transmembrane mechanism (Okumura *et al.*, 2011).

### 2.14.8 Structure and mechanism of parasporins

Parasporins were categorized into two major types based on their amino acid composition, molecular weight, structural homology, and target cell specificity. The three domain parasporins are PS1, PS3, and PS6, whereas the β pore-forming type of parasporins are PS2, PS4, and PS5 (Figure 2.18). The cancer cell killing mechanisms of parasporins were determined to be two different processes: (i) receptor-mediated apoptotic cell death, and (ii) the β pore-forming necrotic cell death mechanism (Akiba & Okumura, 2017).

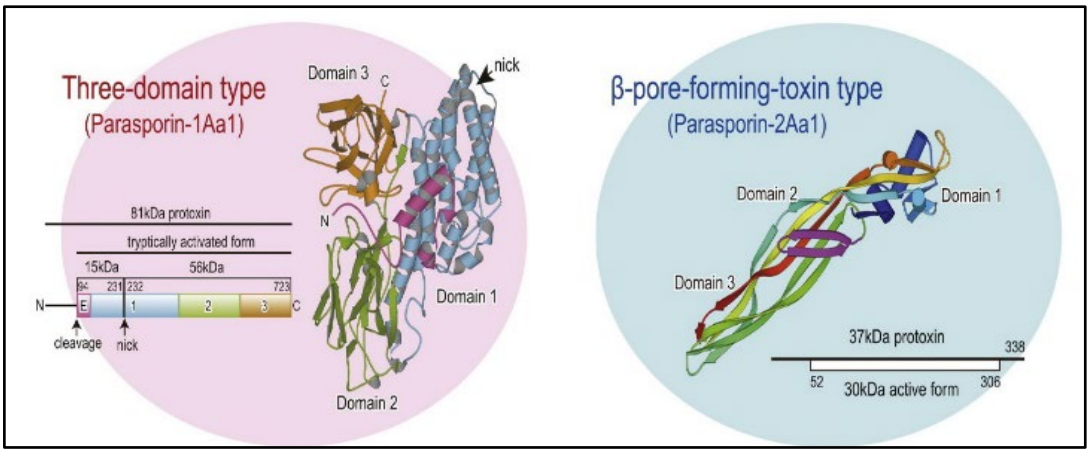



**Figure 2.18:** The three domains and the pore-forming type of parasporins were depicted in the 3D diagram (Akiba & Okumura, 2017).

#### 2.14.8.1 Receptor-mediated cancer cell killing apoptotic mechanism

Parasporin-1 (Cry31), parasporin-3 (Cry41), and parasporin-6 (Cry63) are members of the Cry-type delta endotoxin superfamily of proteins, each consisting of three domains. Researchers have extensively studied parasporin-1 out of these three parasporins. It was shown that parasporin-1 could help Ca2+ ions get into cells from the extracellular buffer without damaging the membrane. The cytotoxicity of parasporin-1 may be suppressed by the G-protein inhibitor suramin, which also inhibits Ca2+ inflow (Katayama *et al.*, 2007) (Figure 2.19). So, it was thought that parasporin-1 would rather cause cell death through apoptosis using a different cytotoxic mechanism than by breaking the membrane(Kitada *et al.*, 2009; Legler *et al.*, 2005; Lingwood *et al.*, 2010). Parasporin-1 had the ability to attach itself to the Beclin-1 receptor in the cell membrane. Beclin-1, a tumor suppressor protein, was discovered as the receptor for parasporin-1 in HeLa cells (Patent No. EP2273266A1, 2011). Furthermore, the scientists also discovered that glyceraldehyde-3-phosphate dehydrogenase (GAPDH) binds to the cell membrane of CEM-SS cells. This makes it easier for the parasporin-like protein of Bt18 to stay in place (Krishnan *et al.,* 2010). Similar to parasporin-1 (Cry31), it was believed that parasporin-3 (Cry41) and parasporin-6 (Cry63), which are also anticancer Cry toxins with three domains, may also kill cancer cells by apoptosis through a receptor-mediated signaling pathway (Figure 2.19).

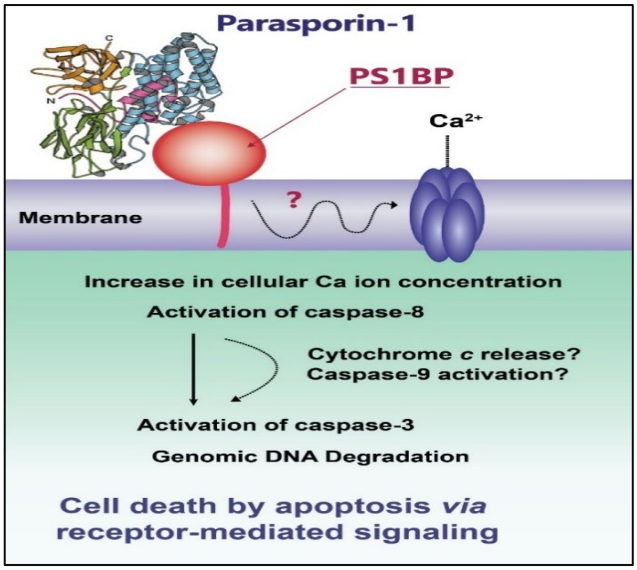



**Figure 2.19:** The diagram illustrates how the three-domain Cry toxin PS1, PS3, and PS6 kill cancer cells through apoptosis via receptor-mediated signaling (Katayama *et al.,* 2007).

### 2.14.8.2 The β pore-forming necrotic cell death mechanism

The β-pore-forming-toxin (β-PFT) type is part of the ETX/MTX superfamily of toxins. Parasporin-2 (Cry46), parasporin-4 (Cry45), and parasporin-5 (Cry64) killed cancer cells through multiple receptors that caused necrotic cell death oligomerization and pore formation (Akiba & Okumura, 2017a). The receptors for parasporin-2 exhibit a little reliance on cholesterol but a clear need for GPI-anchored proteins. The glycan region of the GPI anchor has the potential to facilitate the binding of parasporin-2 to the surface and its subsequent assembly into the membrane (Kitada *et al.*, 2009). As a result, researchers have proposed a model that involves multiple steps: first, parasporin-2 attaches to GPI-anchored receptors or other proteins in the membrane; then, as the toxin concentrates and forms oligomers, transmembrane pores form, leading to impaired membrane permeability (Figure 2.20) (Abe *et al.*, 2017; Kitada *et al.*, 2009).

Another β-pore-forming-toxin (β-PFT) type, the ETX/MTX superfamily toxin protein parasporin-4 (Cry45) was also studied. Parasporin-4 has a nonselective affinity for the plasma membrane, resulting in the formation of a pore complex consisting of several subunits in the cells it targets (Okumura *et al.*, 2011). However, more investigation is required to determine the particular receptor and its transmembrane mechanism (Okumura *et al.*, 2011).

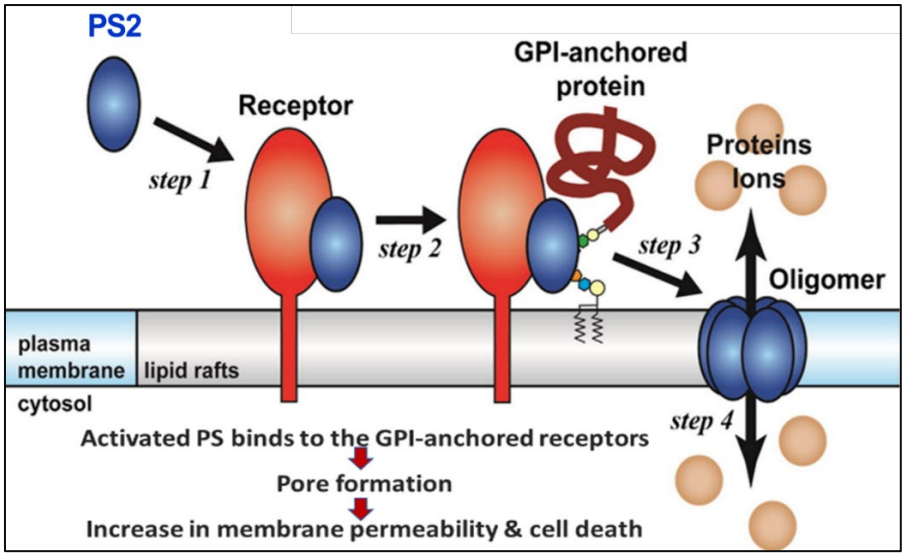



**Figure 2.20:** The diagram illustrates the cancer cell killing mechanism of the β-PFT type PS2. In various cancer cells, necrotic cell death was caused by oligomerizing and pore-forming toxin via overexpressing GPI-anchored protein (Abe *et al.*, 2017).

## 2.15 Surface layer protein

Recently, two strains of *B. thuringiensis*, GP1 (Peña *et al.*, 2006) and CTC (Sun *et al.,* 2001; Zhu & Yu, 2008), were shown to have synthesis parasporal inclusion bodies composed of surface layer proteins (SLP) EA1 and Sap instead of the usual Cry proteins. It is still uncertain if the *sap* and *eag* genes are exclusive to these two *B. thuringiensis* strains or whether they are found elsewhere in *B. thuringiensis*. Recent Western blot tests revealed that 25% of the 59 B. thuringiensis strains examined exhibited a response to antisera specific to the structural component of SLP (Guo *et al.,* 2008). Their findings indicated that 25% of *B. thuringiensis* strains may contain parasporal inclusion bodies made of SLP.

The surface layers, referred to as S-layers, consisted of two-dimensional crystalline arrays of glycoprotein subunits or cell surface proteins (Zhu *et al.,* 2017). These structures were identified as paracrystalline lattices that are formed by the non-covalent assembly of a single (glyco) protein protomer on the surface of the bacterium, driven by entropy. The assembly process described here leads to the formation of intricate and well-organized structures that fully coat the surface of the cell (Sleytr and Sfira, 1986; Sleytr *et al.,* 1993; Sleytr and Beveridge, 1999; Mesnage *et al.,* 2000; Sára and Sleytr, 2000; Pum *et al.,* 2013). Surface layer proteins (SLPs) were found inside the cell

membranes of several types of bacteria and archaea (Figure 2.21). In Gram-negative bacteria, SLPs were often linked with polysaccharides, whereas in Gram-positive bacteria, they were connected with peptidoglycans (Guo *et al.,* 2008). The primary roles of SLP proteins include: (1) engaging with proteins outside the cell; (2) defending against infections; (3) engulfing foreign particles; (4) fortifying membranes; and (5) adhering to surfaces, among other tasks (Sára & Sleytr, 2000). Contrary to the cry genes of *B. thuringiensis*, which were only activated during sporulation, the S-layer genes remain continuously active throughout the whole lifespan of the cell (Guo *et al.,* 2008). The molecular mass of bacterial S-layers ranges from 40 to 170 kDa (Sleytr *et al.,* 1993; Sleytr *et al.,* 2014).

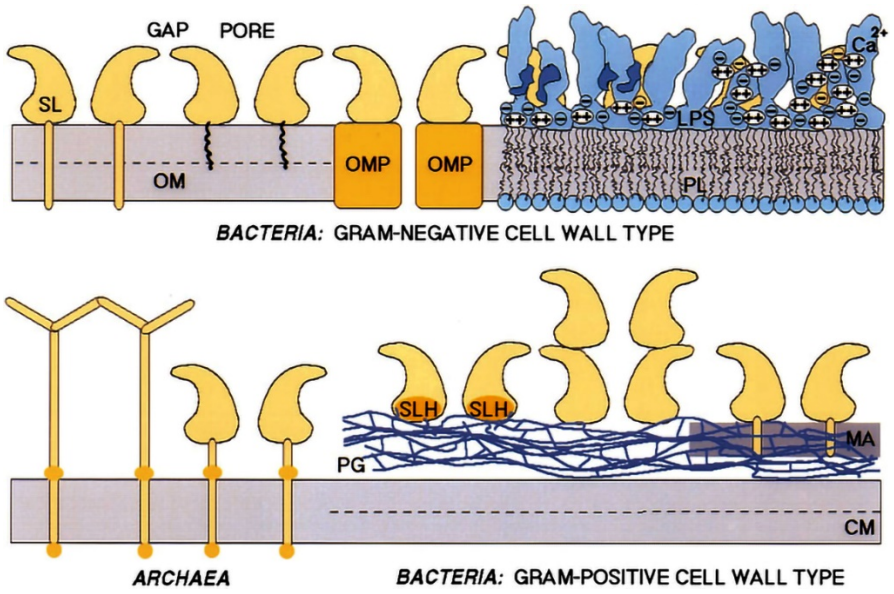



**Figure 2.21:** Scheme of known S-layer-cell envelope interactions with bacteria and archaea. (Top) S-layer (SL) is anchored by a hydrophobic stretch in the outer membrane (OM). (Bottom right) Interactions between S-layer proteins and the peptidoglycan (PG) of Gram-positive bacteria. The (specific) association of a second S-layer is also indicated; double S-layers occur in Gram-positive bacterial membranes (Engelhardt and Peters, 1998).

Two surface layer proteins (SLPs) had been identified in *B. anthracis*: the surface array protein, Sap (Etienne-Toumelin *et al.,* 1995), and the extractable antigen 1, EA1 (Mesnage *et al.,* 1997). The *sap* and *eag* genes, which are controlled throughout development, encode them. According to Mignot *et al.,* (2002), the synthesis of the S-layer Sap occurred during a period of rapid development, while the synthesis of EA1 occurred during a period of stable growth. The exponential phase known as the 'Sap layer' is subsequently replaced by a stationary phase referred to as the 'EA1 layer'. The

regulation of this transition occurs at the transcriptional level. The transcription of *sap* is specifically carried out by RNA polymerase that contains sA, whereas the production of *eag* relies on sH. Sap is essential for the precise regulation of *eag*'s temporal control, whereas EA1 plays a crucial role in tightly regulating *eag* via feedback mechanisms. The regulatory mechanism is likely to be direct since both S-layer proteins have been seen to bind to the *eag* promoter in laboratory conditions. This binding suggests that they may function as transcriptional repressors.

Most of the case *eag* and *sap* genes are located next to each other on the chromosome. Furthermore, they found that the genes are arranged in the same direction, with *eag* positioned downstream from sap. Nevertheless, it is unlikely that the genes are arranged in an operon due to the presence of 722 base pairs of non-coding DNA between them. Furthermore, Engelhardt & Peters (1998) firmly concluded that strains with disrupted sap genes were abundant in EA1. Etienne-Toumelin *et al.,* 1995 identified no promoter consensus sequences in the regulatory domains of sap and other B. anthracis genes.

Both Sap and EA1, which are Gram-positive surface proteins, have three 'S-layer-homology' domain located just after a signal peptide. EA1 and Sap were shown to be co-localized at the cell surface of the bacilli. Nevertheless, EA1 exhibited a higher degree of affinity for the bacterium compared to Sap. Electron microscopy investigations and *in vivo* research conducted with the engineered mutants revealed that EA1 forms the primary framework of the *B. anthracis* S-layer and serves as the predominant antigen associated with the cells (Mesnage *et al.,* 1997).

### 2.15.1 Parasporal surface layer protein of *B. thuringiensis*

*B. thuringiensis* produced a special type of round or oval-shaped parasporal S-layer inclusion protein. In 2001, for the first time, a 100 kDa S-layer protein (SLP) was identified to produce atypical oval-shaped proteinaceous parasporal inclusion bodies in the *B. thuringiensis* strain CTC (Sun *et al.,* 2001). Further advanced study showed that *B. thuringiensis* subsp. *finitimus* strain CTC synthesized round-shaped parasporal inclusion body (Zhu & Yu, 2008) (Figure 2.5 C).

As with the CTC strain, *B. thuringiensis* strain GP1 also produced 100 kDa parasporal S-layer protein (Peña *et al.*, 2006). Similar to delta-endotoxins (Cry and Cyt), SLPs

were also assembled into parasporal positions with round or oval shapes. Zhou *et al.,* (2011) shown through experimentation that *B. thuringiensis* strains CTC, B22, I13, H67, and BMB1152 produce two distinct forms of parasporal SLPs, namely the surface array protein (Sap) and the extractable antigen-1 (EA1). Experimental evidence had demonstrated that the surface array protein coding gene *sap* and the extractable antigen-1 protein coding gene *eag* were either found together in 35% of the analysed *B. thuringiensis* strains or completely lacking in 65% of the strains (Soufiane *et al.,* 2011).

### 2.15.2 Insecticidal and pesticidal S-layer protein of *B. thuringiensis* and other bacteria

The 100 kDa parasporal SLPs of *B. thuringiensis* strain GP1 had found to show selective pesticidal activity against coleopteran beetles (Mexican bean beetle, *Epilachna varivestis*) (Peña *et al.*, 2006). Besides, *B. thuringiensis* strain GP543 had shown selective toxicity against the most harmful ectoparasite cattle ticks (*Rhipicephalus microplus*) of livestock (Lormendez *et al.,* 2019). In addition, the contribution of S-layer proteins of *Lysinibacillus sphaericus* as a mosquitocidal agent was well demonstrated (Allievi *et al.*, 2014; Thanabalu *et al.*, 1991). Besides, the S-layer protein SplA of *Paenibacillus larvae* had been demonstrated as the causative agent of American Foulbrood of honey bees (Poppinga *et al.*, 2012). *Brevibacillus laterosporus* strain UNISS 18 (=NCIMB 41419), had virulence factors against mosquitoes (*Culex pipiens, Aedes aegypti*), the blue bottle fly (*Calliphora vomitoria*), and blowfly (*Lucilia Caesar*) (Bedini *et al.*, 2020). Besides, multiple S-layer proteins of *B. laterosporus* strain UNISS 18 (=NCIMB 41419), had the real virulence factors for mosquitoes, flies, and also against insect house flies (Ruiu, 2023). These studies confirmed that S-layer proteins appear to be part of the toxins and virulence factor complex of this microbial control agent for invertebrate pests. The toxic activity of SLPs is still unclear; it has been suggested that SLPs have a similar insecticidal activity to Cry proteins but with a different mechanism (Peña *et al.*, 2006).

### 2.15.3 Anticancer activity and mechanism of S-layer protein

The majority of the parasporal S-layer proteins (SLPs) in *B. thuringiensis* were discovered to be non-toxic, and the purpose of the gene responsible for coding these SLPs remains uncertain. Recent research found that an 86 kDa parasporal S-layer protein of *B. thuringiensis* strain AP11 had cytotoxic effects on MDA-MB-231 Breast

Cancer Cells, but it did not have any harmful effects on HaCat non-cancer cell lines (Rubio *et al.,* 2017) (Figure 2.22). The scientists proposed that the cytotoxicity of *B. thuringiensis* against MDA-MB-231 may be triggered by the presence of the cadherin-11 (CDH11) cell membrane receptor in MDA-MB-231 cancer cells, which seems to play a role in recognizing the SLP.

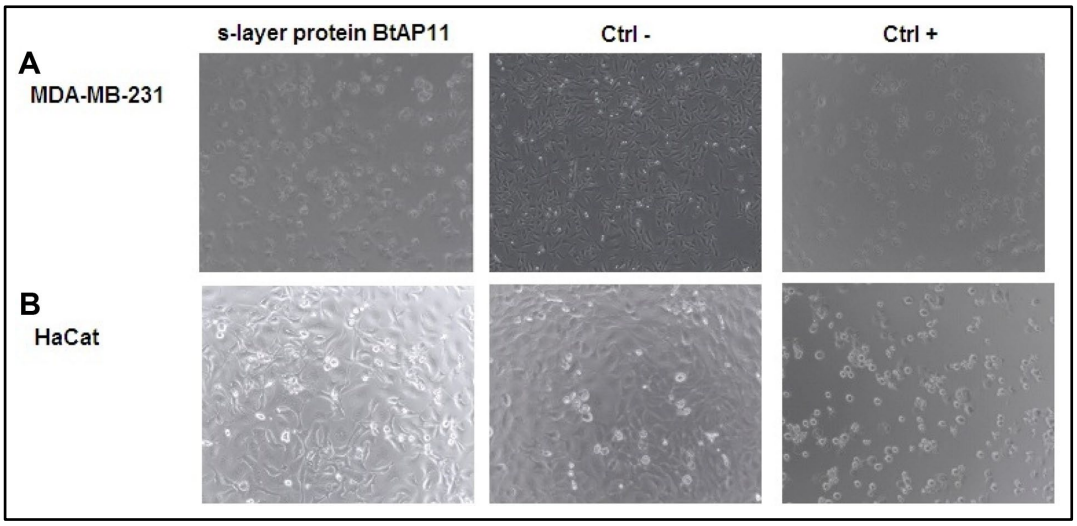



**Figure 2.22:** The cytotoxic effects of the 86 kDa SLP of Bt strain AP11. Toxicities were evaluated in (A) MDA-MB-231 and (B) HaCat cells at a concentration of 0.25 µg/mL. Negative control: cells without SLP; Positive control: cells containing DMSO (Rubio *et al.,* 2017).

Besides, investigation of the antitumor properties of S-layer proteins in lactic acid bacteria (LAB) was also conducted. S-layer proteins in LAB were found to be crucial for adhering to intestinal tissue, along with other functional components (Alp *et al.,* 2020). A recent study had shown that S-layer proteins also contributed to the inhibition of colon cancer, utilizing cell membrane-coated S-CM-HPAD nanoparticles that were boosted by a surface-layer (S-layer) protein. This approach successfully stimulated the immune response against tumors in melanoma models. The findings confirmed that S-layer proteins played a crucial role in suppressing the growth and spread of malignant tumors through the immune system (Wu *et al.,* 2019). The phenomenon of LAB's resilience to cancer is becoming more intriguing and significant. The S-layer protein in LAB had the potential to trigger apoptosis in HT-29 cells via both the death receptor apoptotic route and the mitochondrial pathway (Figure 2.23) (Zhang *et al.*, 2020).

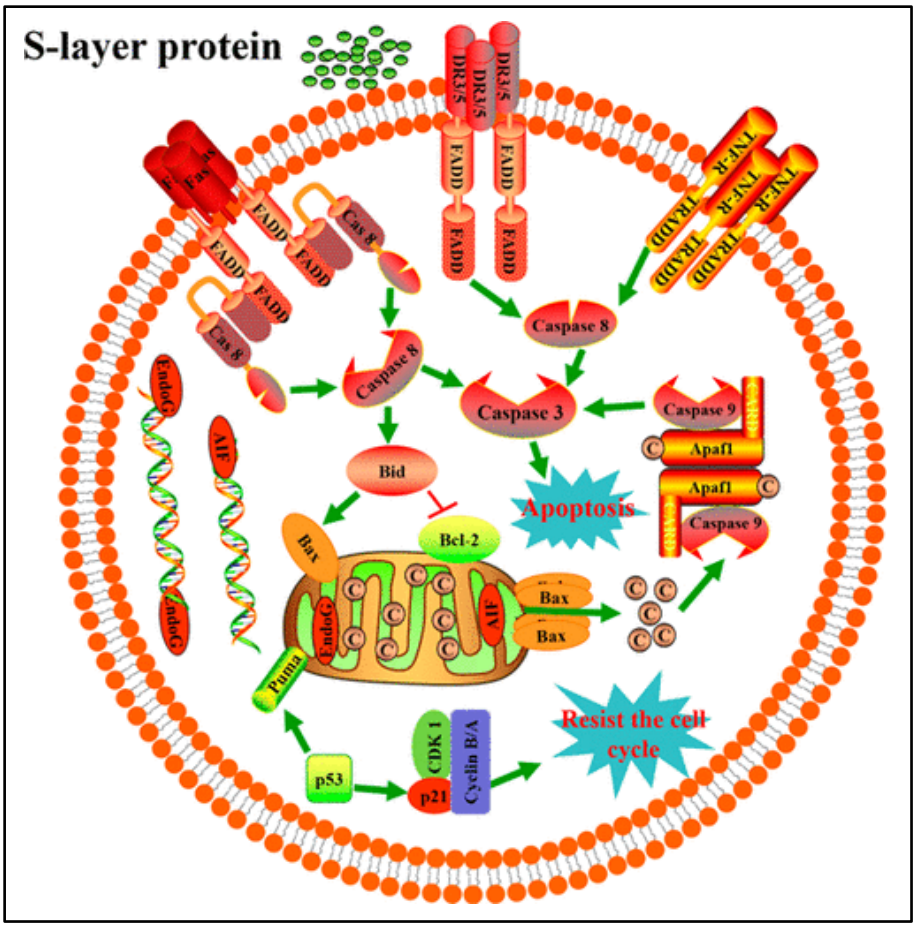



**Figure 2.23:** Illustration of the putative mechanism of action for the S-layer protein. (Zhang *et al.*, 2020).

The S-layer proteins derived from *L. acidophilus* CICC 6074 also exhibited cytotoxic effects on colon cancer HT-29 cells (Liu *et al.,* 2021). This was achieved by impeding the progression of the cell cycle in the G1 phase through the upregulation of p53, p21, and p16 expression, as well as the downregulation of CDK1 (cyclin-dependent kinase) and cyclin B expression (Figure 2.24) (Liu *et al.,* 2021).

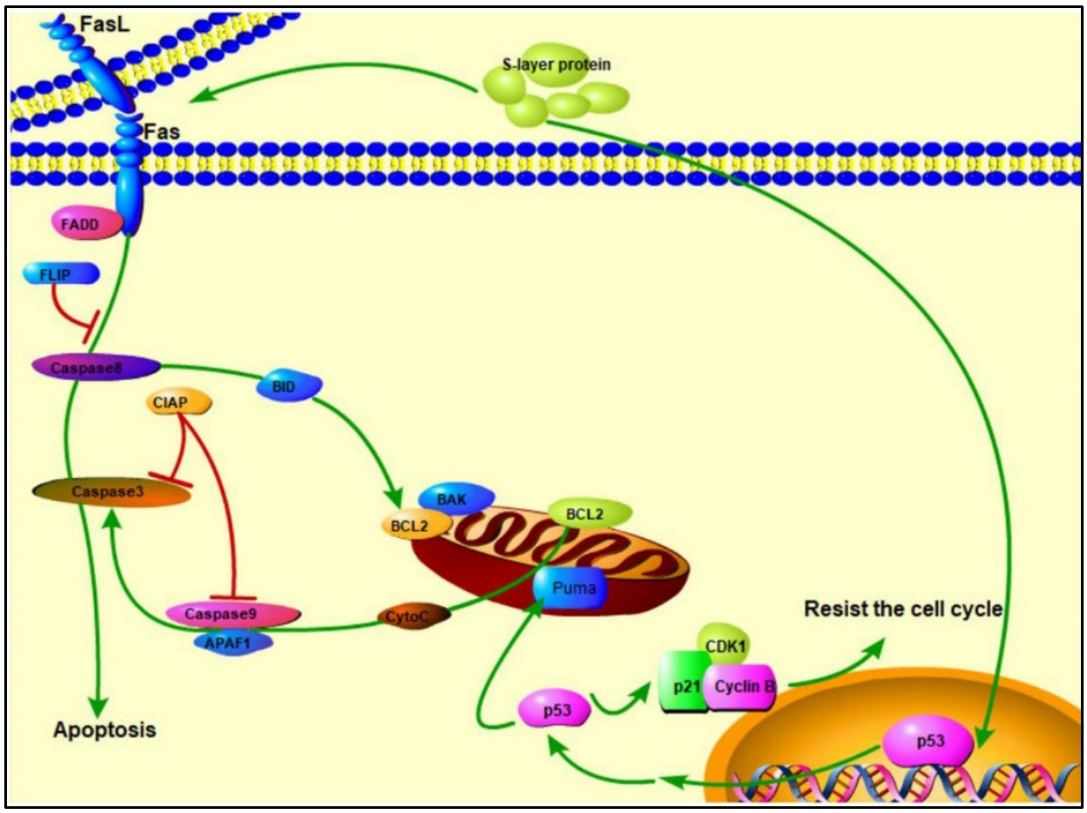



**Figure 2.24:** S-layer proteins of *L. acidophilus* kill cancer cells *via* apoptosis and resist the cell cycle. (→ activation; →→ expression; ⊣ inhibition) (Liu *et al.,* 2021).

### 2.15.4 Domains of bacterial S-layer protein

Electron microscopy and image analysis techniques had shown that all S-layer proteins exhibit two distinct morphological domains (Baumeister & Engelhardt, 1987; Baumeister *et al.,* 1989; Engelhardt & Peters, 1998). Cell-wall-anchoring domain had been identified using biochemical, biophysical, and genetic methods (Jarosch *et al.,* 2001; Lupas *et al.,* 1994; Mesnage *et al.,* 1999; Rünzler *et al.,* 2004; Smit *et al.,* 2002; Smit *et al.,* 2001). Gram-positive bacteria possess diverse anchoring domains that engage in interactions with distinct secondary cell wall polymers. The N-terminal S-layer homology (SLH) domain, consisting of three repeated SLH motifs, was found to interacts with a pyruvylated polysaccharide (Ilk *et al.*, 1999; Mesnage *et al.*, 2000). S-layer proteins lacking SLH domains were found to attach to several secondary cell wall polymers via their N- or C-terminal domains (Egelseer *et al.*, 1998; Jarosch *et al.*, 2000; Sára *et al.*, 1996; Smit *et al.*, 2001).

The second domain of S-layer protein was identified as a crystallization domain (Bingle *et al.,* 1997; Howorka *et al.,* 2000; Jarosch *et al.,* 2001; Mader *et al.,* 2004; Mesnage *et al.,* 1997; Smit *et al.,* 2002). The most recent research had concentrated on examining the connections between the structure and function of this particular area. Mutagenesis techniques, such as point mutations or deletions, had pinpointed specific areas in the crystallization domain that were essential for array formation (Howorka *et al.*, 2000; Jarosch *et al.*, 2001; Sillanpää *et al.*, 2000; Smit *et al.*, 2002). However, these investigations have not yielded a comprehensive understanding of how an S-layer protein undergoes self-assembly to produce a structured arrangement. It was found that only a limited part of the crystallization domain may be removed without compromising its ability to self-assemble (Howorka *et al.*, 2000; Jarosch *et al.*, 2001; Sillanpää *et al.*, 2000; Smit *et al.*, 2002). Limited research groups had conducted structural studies on the crystallization of S-layer proteins (Claus *et al.*, 2002; Jing *et al.*, 2002; Pavkov *et al.*, 2003), and the atomic structure is still unknown. The residues lining the pores and necessary for self-assembly remain to be identified.

The S-layer proteins of Gram-positive bacillus bacteria possessed three consecutives ~55 amino acid repetitions of the (SLH) domain (Fujino *et al.,* 1993; Lupas, 1996; Lupas *et al.,* 1994). Proteins that include three consecutive SLH domains were attached to the bacterial envelope by non-covalent interactions with a secondary cell wall carbohydrate (Mesnage *et al.*, 2000b). The inclusion of chimeric proteins into S-layers was dependent on SLH domains, which weree both essential and sufficient for this process (May *et al.,* 2006; Mesnage *et al.*, 1999). Adjacent to the SLH domains, S-layer proteins included crystallization domains, which were sequences that were expected to facilitate connections between subunits within the S-layer (Candela *et al.,* 2005; Choudhury *et al.*, 2006; Couture-Tosi *et al.*, 2002). As a characteristic of S-layers of bacterial species possessing an outer membrane, usually possess more than one SLH domain with a coiled-coil structure and the β-strand-rich major protein portion (Figure 2.25).

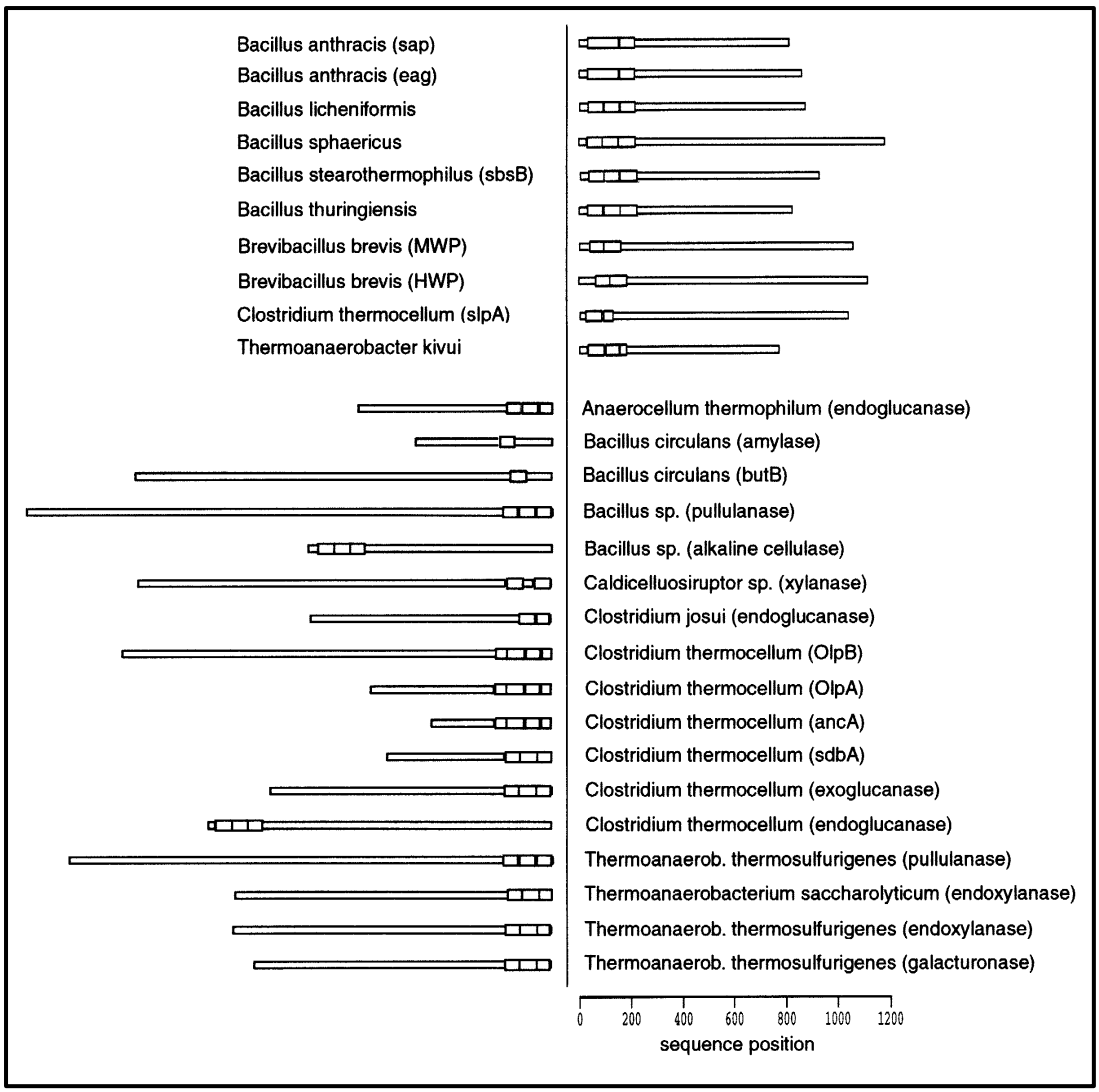



**Figure 2.25:** Organization of SLH domains in the S-layer protein.

Location of SLH domains (white blocks) in S-layers (N-terminal) and extracellular enzymes and functional proteins (C-terminal). White rods either indicate S-layer structural domains (not defined) or several functional domains in extracellular enzymes (not displayed). Note, that the location of SLH domains is consistently close to the N-terminus in S-layers, but mainly at the C-terminal end in extracellular enzymes (Engelhardt & Peters, 1998b).

A recent study on *B. cereus* strain AH187 provided experimental evidence for the presence of two protein components in the S-layer that covers the cell surface of AH187 (Boutonnet *et al.*, 2022). The genetic arrangement of the *B. cereus* AH187 S-layer cluster was revealed, as seen in Figure 2.26. From this experiment it was also found that S-layer protein possessed three consecutive repetitions of the SLH domain. These repeats can bind non-covalently to surface layer homology domains. It was discovered

that the C-terminal region serves as the crystallization domain, enabling the self-assembly of the S-layer proteins (Figure 2.26). The C-terminal crystallization domain of 822-amino acid protein (87.5 kDa) exhibited dissimilarities compared to that of *B. anthracis* SAP, with the two proteins displaying a very modest (33%) level of sequence similarity. By comparison, the C-terminal crystallization domain of 859-amino acid protein (91.0 kDa) and *B. anthracis* EA1 exhibited a high degree of similarity, with a sequence similarity of 89%. Considering these attributes, 822-amino acid protein (87.5 kDa) and 859-amino acid protein (91.0 kDa) and were rebranded as SL2 (representing S-layer 2 protein) and EA1, correspondingly. The genes responsible for their encoding were *sl2* and *eag,* respectively (Boutonnet *et al.*, 2022).

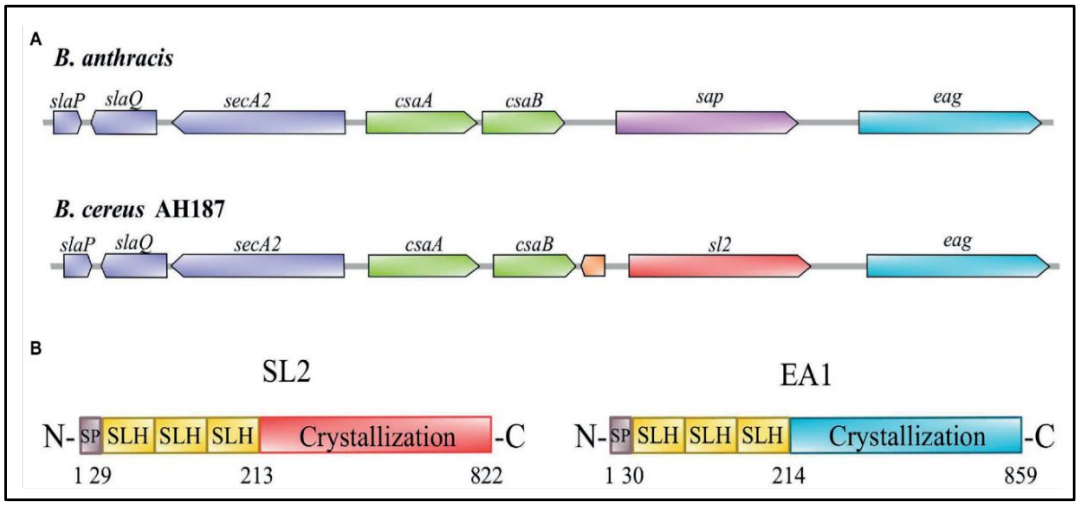



**Figure 2.26:** Locus and main structures of the S-layer proteins from *B. cereus* AH187.

(A) The S-layer locus of *B. anthracis* is being compared to the locus of B. cereus AH187. An unidentified protein is encoded by the open reading frame (ORF) situated between *csaB* and *sl2* in the *B. cereus* S-layer locus. (B) Depiction of the main structures of SL2 and EA1 symbolically or metaphorically. The proteins are composed of three primary components: a signal peptide (SP, gray), three S-layer homology domains (SLH, yellow), and a crystallization domain (red for SL2; blue for EA1) (Boutonnet *et al.*, 2022).

### 2.15.5 3D crystallographic structure of bacterial S-layer protein domains

From x-ray crystallography, the structure of three SLH domains of S-layer protein was assumed to be the shape of a three-prong spindle (Figure 2.27) (Kern *et al.*, 2011). Every SLH domain contributed to a three-helical bundle located at the base of the spindle, while an additional α-helix and its connecting loops formed the three prongs. The inter-

prong grooves of SLH domains contained conserved cationic and anionic residues, which were essential for binding the *B. anthracis* SCWP. Modeled findings indicated that the SLH domains of other S-layer proteins similarly adopted a three-prong spindle shape and similarly bound bacterial carbohydrate envelopes.

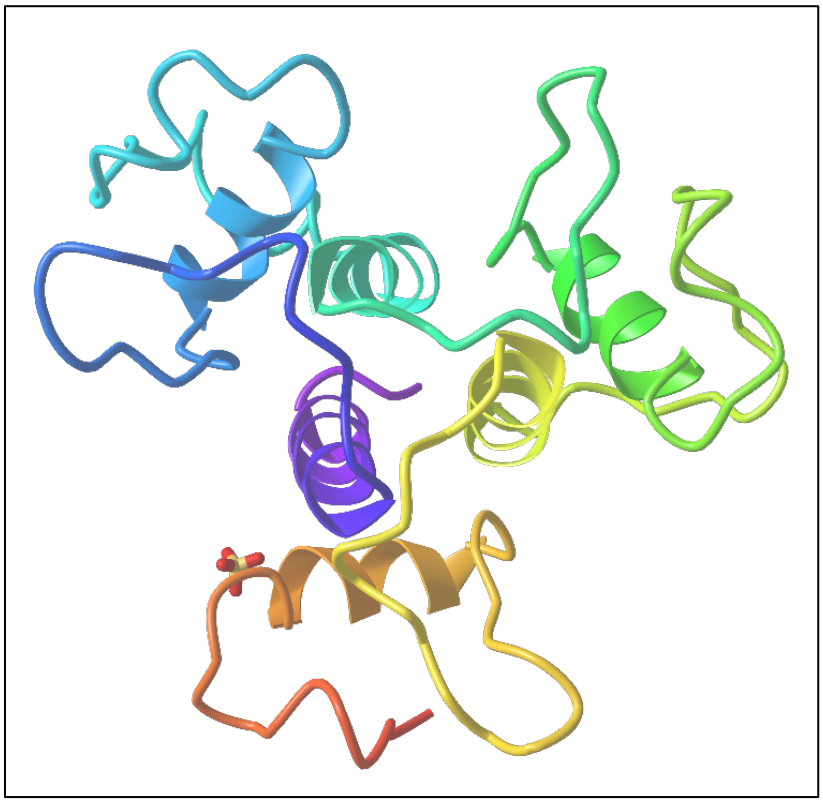


**Figure 2.27:** The 3D view of the crystallographic structure of SLH domains of surface array protein Sap. Source: RCSB PDB, id: 3PYW (Kern *et al.,* 2011).

Fioravanti *et al.*, (2019) reported that Sap consisted of roughly 800 amino acids, which showed a high level of similarity in *B. anthracis*, *Bacillus cereus*, and *B. thuringiensis*. The various strains of these bacteria showed an average 80% sequence identity. An N-terminal signal peptide directed the Sap protein to attach itself to a specific unit in the peptidoglycan known as ketal-pyruvylated N-acetylmannosamine on the cell surface. This binding was facilitated by an α-helical domain that was part of the cell wall and consists of three S-layer homology (SLH) regions (Figure 2.28 A) (Kern *et al.*, 2011; Sychantha *et al.*, 2018). The projected S-layer assembly domain ($Sap^{AD}$) spanned residues 216 to 814 and consisted of six β-sandwich domains (D1 to D6) connected by short linkers (Fioravanti *et al.*, 2019) (Figure 2.28 A, B).

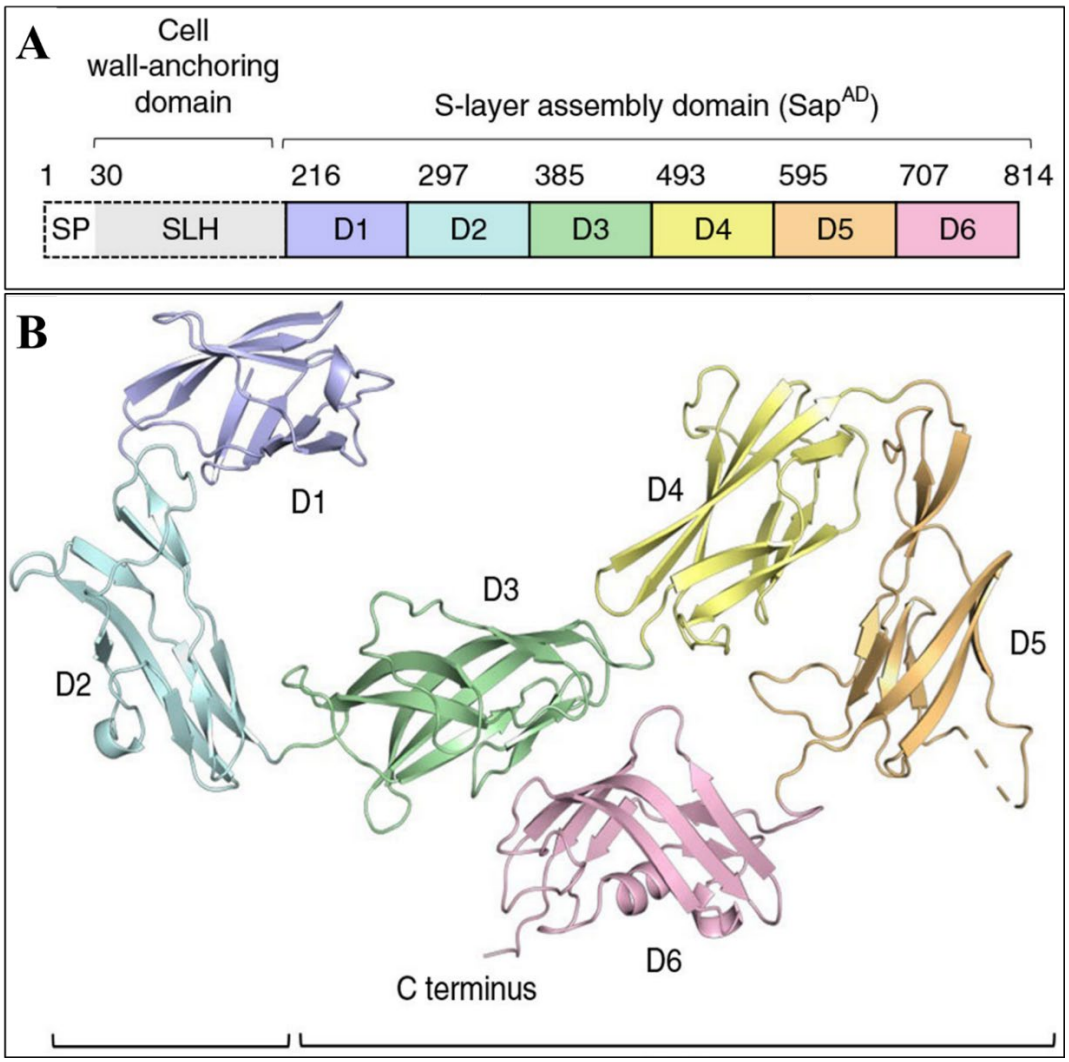



**Figure 2.28:** The X-ray structure of *B. anthracis* Sap assembly domain.

(A) The domain layout of Sap, where the N-terminal first ~215 residues consisting of a signal peptide (SP) and a pseudo repeat of three SLH domains that create a cell wall-anchoring domain, and the Sap assembly area ($Sap^{AD}$) consists of six distinct domains (D1 to D6); (B) Residues 216–814 is shown with a ribbon model, which consists of six distinct β-domains (D1 to D6). These domains come together to form a flat, tile-like unit (Fioravanti *et al.*, 2019).

## 2.16 Aims and Objectives of the Study

Cancer ranks as the second most prevalent cause of mortality. However, the mortality rate from cancer has decreased during the past two decades. Advancements in early detection and cutting-edge therapies are effectively eradicating cancer and significantly extending the lifespan of individuals affected by the disease. New types and natures of cancer are constantly being detected, which is a major challenge for cancer treatment. Because of the diversified cells in the human body, different types of cancer require different drugs to treat them. One of the best features of parasporins of *B. thuringiensis* is their tissue-specific anticancer activity. Besides, the selective anticancer activity of surface layer proteins also began to be explored. Therefore, these types of parasporal proteins with tissue-specific cancer cell-killing abilities could be investigated as possible new and suitable protein based therapeutic candidates for cancer treatment. For this purpose, we aimed to investigate a collection of previously isolated, identified and characterized indigenous *B. thuringiensis* strains (Shishir *et al.*, 2014) for the presence of anticancer parasporal proteins. Therefore, the specific aims and objectives of the study are:

1. Isolation and identification of hemolytic and nonhemolytic *B. thuringiensis* strains from previously isolated two hundred ninety-one (291) indigenous strains of *B. thuringiensis*.
2. Molecular identification of parasporin (*ps*) genes from indigenous *B. thuringiensis* strains by polymerase chain reaction (PCR).
3. Screening of parasporal anticancer protein-producing nonhemolytic and hemolytic *B. thuringiensis* strains.
4. Production of parasporal protein and their extraction, solubilization, purification, and SDS-PAGE analysis.
5. Cytotoxicity determination of solubilized parasporal (Cry/SLP) protein of *B. thuringiensis* strains on brine shrimp (*Artemia* nauplii) and finding out insecticidal activity on Tephritidae fruit fly (*Bactrocera cucurbitae*) larvae.
6. Morphological study of parasporal inclusions of *B. thuringiensis* strains by scanning electronic microscope and their homology analysis with Cry or SLP of other Bt strains.
7. Examination of *in vitro* cytotoxicity of activated parasporal (Cry/SLP) proteins of indigenous *B. thuringiensis* strains on carcinoma cell line (HeLa) and non-carcinoma Vero cell line.
8. Whole genome sequencing (WGS) of anticancer parasporal protein-synthesizing *B. thuringiensis* strains and data analysis for identification, and molecular characterization of parasporal anticancer (Cry/SLP) proteins.

# CHAPTER 3:

# Molecular identification of parasporin genes from indigenous *Bacillus thuringiensis* strains

# CHAPTER 3: Molecular identification of parasporin genes from indigenous *Bacillus thuringiensis* strains

## 3.1 Introduction

*Bacillus thuringiensis* (Bt) is well-known in medical, veterinary, and agricultural microbiology (Travis & Glare, 20000). Parasporal crystalline inclusion proteins of some Bt are of paramount importance due to their insecticidal, nematocidal, and cancer cell-killing capabilities. The genes that make parasporal crystalline inclusion are called Cry genes and are turned on during sporulation because the RNA polymerase that controls Cry gene expression is made only when spores are being made. Cancer cell-killing Cry protein synthesizing genes are called parasporins (*ps*). Like the insecticidal Cry gene, parasporins are also genealogically heterogeneous because of the very low-level (<25%) nucleotide and amino acid sequence homologies found between existing parasporin (PS) and the existing classes of Cry and Cyt proteins (Mizuki *et al.*, 2000). Parasporins are very rare in the Bt, and since its first discovery in 2000, only 19 parasporins have been identified (https://www.fitc.pref.fukuoka.jp/parasporin/list.html).

Both hemolytic and nonhemolytic Bt strains could harbor the parasporin genes (Ichikawa *et al.*, 2008). To screen and identify cry genes at the molecular level from a large size Bt sample, traditional PCR was the faster and low-cost approach where a set of universal primers were used (Asokan *et al.*, 2014; Ben-dov *et al.*, 1997; Ben-Dov *et al.*, 1999; Ceron *et al.*, 1994, 1995; Porcar et Juárez-Pérez, 2003).

Although parasporins are genealogically heterogeneous, an attempt was made to identify them through PCR from indigenous non-insecticidal Bt strains. Initially, screening was done to separate nonhemolytic Bt from the hemolytic strains among the previously isolated Bt collection of Enzyme & Fermentation Laboratory, Department of Microbiology, University of Dhaka. In searching novel parasporins from indigenous Bt, a total of 98 nonhemolytic and previously identified 23 non-insecticidal hemolytic Bt strains were screened. Besides, 3 reference insecticidal strains, HD-73, sotto, and japonensis were also studied. Parasporin gene-specific (*ps1, ps2, ps3, ps4,* and *ps5)* primers were constructed using the conserved sequence of the parasporins found in the NCBI database. To identify parasporins, these newly designed primers were employed, and both plasmid as well as genomic DNA of the indigenous non-insecticidal *B.*

*thuringiensis* strains were used as templates. Parasporins primer specific amplified PCR amplicons were then subjected to DNA sequencing followed by data analysis.

## 3.2 Materials and Methods

### 3.2.1 Materials

#### 3.2.1.1 Chemicals

Agar ($C_{14}H_{24}O_9$), acetic acid ($CH_3COOH$), boric acid ($H_3BO_3$), $dH_2O$, $Na_2EDTA$ ($C_{10}H_{14}N_2O_8.2Na.2H_2O$), ethanol ($CH_3CH_2OH$), ethidium bromide ($C_{21}H_{20}BrN_3$), glycerol ($C_3H_8O_3$), HCl, methanol ($CH_3OH$), $MgCl_2$, dNTP, Taq polymerase, KCl, NaCl, $Na_2HPO_4$, $KH_2PO_4$, NaOH, peptone, tris-base ($C_4H_{11}NO_3$). The chemicals used in this study were of molecular grade, and their lists and company names are mentioned in **Appendix III**.

#### 3.2.1.2 Types of equipment

Autoclave machine, biosafety cabinet, centrifuge machine, electronic balance, glassware sterilizer, microbiological incubator, magnetic stirrer, micropipettes, nanodrop 2000, orbital shaker incubator, pH meter, agarose gel electrophoresis power supply, -20 $^oC$ fridge, 4 $^oC$ refrigerator, spectrophotometer, thermal cycler, thermo stated shaking water bath, vortex mixture, etc., apparatuses used in this study and were mentioned in the respective methods section. Their company and models were mentioned in **Appendix IV**.

#### 3.2.1.3 Media

Luria Bertani (LB) agar and broth, nutrient agar (NA), and sheep blood agar media used in this study and their compositions are mentioned in **Appendix I**.

#### 3.2.1.4 Kits, reagents and solutions

Phosphate-buffered saline (PBS), normal saline, plasmid isolation kit (Favorgen), 2X PCR master mix (Invitrogen), PCR purification kit (PureLink™, Invitrogen) TAE, TBE, TE buffer, Tris-HCl, EDTA, etc., reagents and solutions used in these experimental procedures, and their composition and preparation methods are mentioned in **Appendix II.**

### 3.2.1.5 Bacterial strains

This research was carried out by selecting and screening 291 indigenous *B. thuringiensis* strains previously isolated from different locations and ecosystems in Bangladesh. Besides, 3 strains, kurstaki HD-73, sotto, and japonensis, which were acquired from the Bt stock collection of Okayama University in Japan, were utilized as reference strains. For this research, two different kinds of *Bt* strains were investigated. A total of ninety-eight nonhemolytic Bt strains (Table 3.2.1) and twenty-three hemolytic non-insecticidal *B. thuringiensis* strains were studied (Table 3.2. 2).

**Table 3.2.1:** List of 98 nonhemolytic *B. thuringiensis* strains studied

| Serial | Name of Bt strains | Serial | Name of Bt strains | Serial | Name of Bt strains | Serial | Name of Bt strains |
|---|---|---|---|---|---|---|---|
| 1. | BySa2 | 26. | FhSd2 | 51. | NaSc2 | 76. | SgSm1 |
| 2. | CgSb1 | 27. | JeSe1 | 52. | NaSc3 | 77. | SgSn2 |
| 3. | CgSc2 | 28. | KbSb2 | 53. | NaSd1 | 78. | SgSp2 |
| 4. | ChSa1 | 29. | KbSc1 | 54. | Nasd2 | 79. | SpSa1 |
| 5. | ChSd2 | 30. | KbSc2 | 55. | NsSc2 | 80. | SpSa2 |
| 6. | Dda3 | 31. | KfSa2 | 56. | RaSa1 | 81. | SpSb2 |
| 7. | DpSb1 | 32. | KfSb2 | 57. | RaSb1 | 82. | SpSb3 |
| 8. | Dsa3 | 33. | KkSb1 | 58. | RaSc1 | 83. | SpSc2 |
| 9. | Dsd3 | 34. | KkSb2 | 59. | RaSd1 | 84. | SpSc3 |
| 10. | Dse1 | 35. | KkSc1 | 60. | RaSe1 | 85. | SpSd1 |
| 11. | Dsf1 | 36. | KkSd1 | 61. | RfSa2 | 86. | SpSd3 |
| 12. | Dsf7 | 37. | Ksa2 | 62. | RfSc3 | 87. | SpSd3 |
| 13. | Dsf8 | 38. | Ksb1 | 63. | RhSa2 | 88. | SpSe2 |
| 14. | Dsh1 | 39. | Ksc1 | 64. | RhSa2 | 89. | Ssa1 |
| 15. | Dsh3 | 40. | MaSb1 | 65. | RhSb3 | 90. | Ssa3 |
| 16. | Dsh5 | 41. | MeSb3 | 66. | RhSc3 | 91. | Ssd1 |
| 17. | Dsi1 | 42. | MeSd1 | 67. | RhSc4 | 92. | Ssd2 |
| 18. | Dsi2 | 43. | MeSd2 | 68. | RhSd1 | 93. | TaSa1 |
| 19. | Dsn5 | 44. | MuSd1 | 69. | RhSd3 | 94. | TaSa3 |
| 20. | FHSa1 | 45. | MyIa1 | 70. | RpSb2 | 95. | TaSb3 |
| 21. | Fhsb2 | 46. | MyIa2 | 71. | SaSb1 | 96. | TaSc2 |
| 22. | FhSb3 | 47. | MyIb1 | 72. | SaSc1 | 97. | TaSd1 |
| 23. | Fhsc1 | 48. | MySb1 | 73. | Sasd1 | 98. | TaSe2 |
| 24. | FhSc4 | 49. | NaSb2 | 74. | SaSe1 | | |
| 25. | FhSd1 | 50. | NaSc1 | 75. | SgSj2 | | |

**Table 3.2.2:** List of 23 indigenous and 3 reference hemolytic *B. thuringiensis* strains studied

| Serial | Name of Bt strains |
|---|---|
| 1. | 1i |
| 2. | 19S |
| 3. | 25L |
| 4. | 28S |
| 5. | 34L |
| 6. | 45L |
| 7. | 55S |
| 8. | 57S |
| 9. | BD59S |
| 10. | Leaf-28 |
| 11. | Leaf-31 |
| 12. | Leaf-58 |
| 13. | Soil-108 |
| 14. | Soil-46 |
| 15. | Soil-57 |
| 16. s | Dse1 |
| 17. | Dsf7 |
| 18. | Dsh4 |
| 19. | DSN4 |
| 20. | DSN7 |
| 21. | JeSa1 |
| 22. | JSa3 |
| 23. | SSa3 |
| 24. | Reference Bt strain sotto |
| 25. | Reference Bt strain HD-73 |
| 26. | Reference Bt strain japonensis |

## 3.2.2 Methods

### 3.2.2.1 Hemolytic activity test

For separation of hemolytic and nonhemolytic *B. thuringiensis*, previously isolated two hundred ninety-one (291) native strains were examined on sheep blood agar plates. The red blood cells in blood agar serve as a solid growth substrate. Hemolysis is a process whereby bacteria break apart blood cells, and this medium is used to identify bacteria that generate such enzymes. Bacteria were differentiated from one another based on the extent to which they hemolyzed red blood cells. Sheep erythrocytes were separated from whole blood by adding sterile glass beads, and after continuous shaking, red blood

cells were separated. Red blood cell containing blood serum was added to the blood agar. Hemolytic activity was determined by placing a single bacterial spot with a needle onto blood agar containing 5–8% sheep blood. The plate was then incubated in an incubator at 37 °C for 16–18 hours. Bt strains were classified according to their hydrolysis pattern: beta-hemolytic, by the presence of large clear zones surrounding the colonies; alpha-hemolytic, a greenish halo surrounding the colonies; and nonhemolytic, without any yellowish browning or greenish halo surrounding the colonies.

### 3.2.2.2 Primer designing

Six sets of primer pairs (Table 3.2.3) were designed from *ps1* through *ps5* genes (*ps1Con*, *ps2*, *ps3N*, *ps3CryBP*, *ps4*, and *ps5*) respectively from conserved nucleotide blocks found in the δ-endotoxin-N, CryBP1, β-pore-forming toxin, and ETX-MTX toxin coding regions using FastPCR (Kalendar *et al.,* 2017; Kalendar *et al.,* 2014), and PerlPrimer v1.1.21 (Marshall, 2004) software. A diagram of the conserved toxin region is shown in Figure 3.2.1. Then, primer qualities (specificity, primer dimer, PCR quality, linguistic complexity, and so on) were assessed by FastPCR (PrimerDigital), PerlPrimer, and OligoAnalyzer3.1 (Integrated DNA Technologies). Then binding specificity of each primer to each gene was checked by local alignment using MEGA6 (Tamura *et al.,* 2013) and NCBI primer BLAST (Johnson *et al.*, 2008) search tool. All primers were prepared according to the synthesizer's (Macrogen, Korea) instructions, dissolved in $dH_2O$ and stored in a -20 °C fridge.

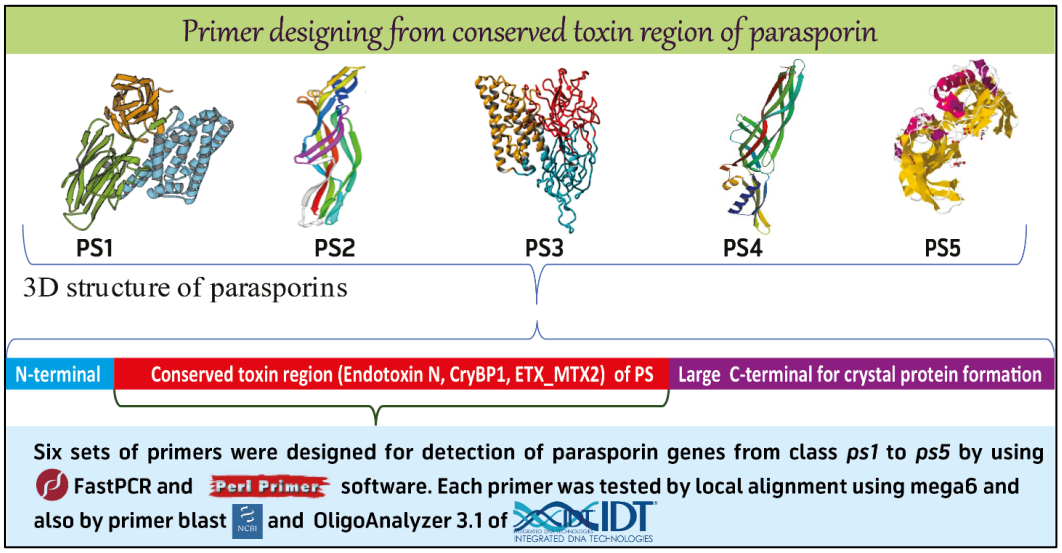



**Figure 3.2.1:** Diagram showing the conserved toxin region of parasporins used for primer designing.

**Table 3.2.3:** List of primers used for detection of *ps* genes of indigenous *Bt* strains

| Primer ID | Sequence (5'-3') | Length | CG (%) | Tm (°C) | Linguistic complexity | Primer Quality | PCR Fragment Size(bp) | NCBI BLAST |
|---|---|---|---|---|---|---|---|---|
| PS1Con-F | TTATGGGAGCAGAACTCGCA | 20 | 50 | 55.9 | 83% | 71% | 542 | 100% |
| PS1Con-R | AAAGCACTCATAGCCATTCC | 20 | 45 | 53.4 | 95% | 72% | | 100% |
| PS2-F | GTTGGGACTGTTCAGTACGT | 20 | 50 | 54.9 | 86% | 88 | 509 | 100% |
| PS2-R | AGATACATTCGTCACTGTACCTC | 23 | 43.5 | 53.8 | 90% | 79 | | 100% |
| PS3N-F | GAGAACCGAATGGTAGGAAGTC | 22 | 50 | 55.7 | 85% | 70% | 258 | 100% |
| PS3N-R | CTGTTGTGTCCGTGCATAGTA | 21 | 47.6 | 54.9 | 82% | 82% | | 100% |
| PS3BP-F | AATGGAAATTACCCACCGCT | 20 | 45 | 54.5 | 86% | 89% | 499 | 100% |
| PS3BP-R | ACCGAAGAAACTGACTTCCT | 20 | 45 | 53.3 | 78% | 82% | | 100% |
| PS4-F | TACGCCAAATCCAACTAGTCAG | 22 | 45.5 | 54.9 | 85% | 80% | 388 | 100% |
| PS4-R | CTATTCCAGTACCAGTGTAACCAG | 24 | 45.8 | 55.3 | 74% | 74% | | 100% |
| PS5-F | ACTGAATTTAGTTCAACGCCAC | 22 | 40.9 | 54.1 | 98% | 87% | 616 | 100% |
| PS5-R | GCAGTTATTCTAGCCTCACC | 20 | 50 | 53.8 | 87% | 80% | | 100% |

### 3.2.2.3 Genomic DNA extraction and preservation

Genomic DNA extraction from *B. thuringiensis* strains was performed following the method described by (Bravo *et al.*, 1998). In order to get genomic DNA from the *B. thuringiensis* strain, a single colony was inoculated on an LB agar medium during the process. After overnight incubation, a single colony was transferred into a microcentrifuge tube (1.5 mL) and mixed well with 100 µL of sterile distilled water by vortex. This bacterial suspension was then placed immediately into boiling water for 10 minutes to lyse the cells and allow the chromosomal and plasmid DNA to come out of the cells. The resulting cell lysate was then briefly centrifuged at 11,000 × g for 8 min, and the concentration of the supernatant was measured using NanoDrop 2000 (Thermo Fisher Scientific). Supernatants containing DNA with a good concentration (250–400 ng/µL) were collected in a new, fresh tube, stored at -20 ºC, and were used as template DNA for PCR.

### 3.2.2.4 Plasmid DNA extraction and preservation

Bt strains were grown in LB broth, and plasmid extraction was done using the Favorgen Plasmid Extraction Mini Kit following the manufacturer's instructions. In summary, the *B. thuringiensis* strains were cultivated in LB broth. A thriving bacterial culture (about 1-3 mL) was then transferred to a centrifuge tube. In order to collect the cells, the tubes underwent centrifugation at 11,000 × g for 1 minute. The cell pellet was resuspended in 200 µL FAPD1 buffer (with RNase A added) by pipetting, followed by the addition of 200 µL FAPD2 buffer. Samples were then incubated at room temperature for 5

minutes to allow the cells to lyse. Then 300 µL of FAPD3 buffer was added for neutralization of the lysate, which was then centrifuged at 18,000 × g for 5 minutes to clarify the lysate. The supernatant was transferred to the FAPD column, then placed in a collecting tube, and centrifuged at 11,000 × g for 30 seconds. Then, the flow-through was discarded, and 400 µL of W1 buffer was added to the FAPD column, put in the collecting tube and centrifuged at 11,000 × g for 30 seconds. The column was put back in the collecting tube, and the flow-through was thrown away. Then 700 µL of wash buffer was added to the FAPD column and was centrifuged at 11,000 × g for 30 seconds. The flow-through was then discarded, and the column was put back in the collecting tube to centrifuge at maximum speed (18,000 × g) for 3 minutes to dry the FAPD column. Finally, the FAPD column was placed in a sterilized 1.5-ml microcentrifuge tube, and 50 µL of elution buffer was added to the center of the FAPD column. After waiting for 1 minute, centrifugation was done at maximum speed (18,000 × g) for 1 minute to collect plasmid DNA. The purified plasmid can be stored at -20 °C for further use.

#### 3.2.2.5 Quantification of DNA

NanoDrop 2000 was used to analyze DNA concentration (Thermo Scientific). One µL of nuclease-free water or plasmid elution buffer was used as blanks. Then DNA concentration (ng/µL) of 1 µL of the sample was measured at $OD_{260}/0D_{280}$. The $OD_{260}/OD_{280}$ ratio for a pure DNA preparation is 1.8 (Sambrook, 1982).

#### 3.2.2.6 Screening of parasporin genes by polymerase chain reaction (PCR)

Both plasmid and genomic DNA of each of the 124 Bt strains were examined to amplify the targeted gene sequence of *ps-1, ps-2, ps-3, ps-4,* and *ps-5* using each pair of newly designed primers. The PCR reactions were performed in a thermal cycler (ProFlex™ 3 x 32-well PCR System) with a reaction mixture containing 10 µL of 2X PCR master mix (Invitrogen) and 1 µL of genomic DNA (250–300 ng) or plasmid DNA (30 ng) of indigenous Bt strains, along with 1 µL (10 picomoles) forward and 1 µL (10 picomoles) reverse primer, and 7 µL nuclease-free water for a final volume of 20 µL (Table 3.2.4). The PCR program was started with a denaturation step followed by a three steps amplification for 35 cycles and a single final extension step (Table 3.2.5). The differences in annealing temperature depend on the use of varied primers for the

amplification of different *ps* genes. In brief, template DNA was preheated at 95 °C for 5 minutes, followed by 35 cycles of denatured (95 °C for 45 seconds), annealed (gene-specific temperature for 1 minute) and extensions (72 °C for 1 minute) to achieve the PCR products. Then, a single final extension was done at 72 °C for 10 minutes.

**Table 3.2.4:** Ingredients and quantity of reagents for a 20 µL PCR reaction

| Ingredients | Volume Added |
|---|---|
| 2X PCR master mix ($MgCl_2$, dNTP, Taq Polymerase) | 10 µL |
| Forward primer | 1 µL (10 picomole/µL) |
| Reverse primer | 1 µL (10 picomole/µL) |
| Template | 1 µL (250 -350 ng genomic DNA or 30 ng plasmid DNA) |
| PCR Grade Water | 7 µL |
| Final volume | 20 µL |

**Table 3.2.5:** PCR reaction condition

| Steps | State of PCR | Temp (ºC) | Time |
|---|---|---|---|
| **Step-1** | The first stage of denaturation/ preheat | 95 | 5 minutes |
| **Step-2** | Denaturation | 95 | 45 seconds |
| | *Annealing | 53-56 | 1 minute |
| | Extension | 72 | 1 minute |
| **Step-3** | Final extension | 72 | 10 minutes |

*(*ps1Con*: 53 ºC, *ps2*: 55 °C, *ps3N*: 56 °C, *ps3CryBP*: 55 °C, *ps4*: 56 °C, *ps5*: 55 °C)

### 3.2.2.7 Agarose gel electrophoresis and observation of the amplified PCR bands

The resulting PCR products were separated using agarose gel electrophoresis (Sambrook *et al.,* 1989). A 1% agarose gel was made by combining 0.5 grams of agarose with 50 milliliters of 1X Tris-acetate-EDTA buffer. Subsequently, the recipe was subjected to microwave irradiation for approximately 2.0-2.5 minutes to ensure full

dissolution of the agarose. The liquid gel was supplemented with ethidium bromide at a final concentration of 0.5 µg/mL and thereafter poured over the tray. A comb of the specified dimensions was placed in the tray and allowed to cool at room temperature until the gel solidified. After the agarose gel had set and was ready for use, the comb was cautiously removed to ensure that the wells remained intact. Subsequently, the gel was immersed in the electrophoresis tank containing 1X TBE buffer. 5 µL of the PCR product that underwent amplification was combined with 2-3 µL of gel loading dye, specifically the 6X MBI Fermentas variant, and thereafter placed into each individual well of the gel. The electrophoresis was conducted at a voltage of 80 volts (Bio-Rad, USA). Additionally, 5 µL of a 100 bp plus marker DNA (Invitrogen, USA) was included in the run alongside the samples to confirm the amplification of the required gene with the precise size of the PCR product. The gel was subsequently seen under UV trans-illumination using a gel documentation system (Gel Doc, Bio-Rad, USA). The images were acquired using a Gel Doc imaging system (Bio-Rad, USA) linked to a computer. The bands were analyzed using 'Alpha Imager' software (Bio-Rad, USA).

#### 3.2.2.8 Purification of PCR Products

Amplified PCR products were purified by using the PureLink™ PCR Purification Kit (Invitrogen) following the manufacturer's instructions. In brief, approximately 100 µL PCR products were transferred into a 1.5 mL microcentrifuge tube. Then, 400 µL of binding buffer was combined with isopropanol and well-mixed. Subsequently, the sample containing the binding buffer was introduced into the PureLink spin column. The column was then centrifuged at room temperature with a force of 10,000 × g for 1 minute. The pulled flow-through was discarded and the column was again inserted into the collecting tube. The column was enriched with 650 µL of Wash Buffer and centrifuged at 10,000 × g for 1 minute. The flow-through was discarded. The column was then centrifuged at 10,000 × g for 2-3 minutes. The collecting tube was disposed of. Then the spin column was inserted into a sterile 1.7 mL PureLink collection tube provided with the package. A volume of 50 µL of elution buffer, consisting of 10 mM Tris-HCl at a pH of 8.5, was introduced into the central region of the column. Then the

column was incubated at ambient temperature for 1 minute. Then, the column was centrifuged at 15800 × g for 2 minutes. Purified PCR products were collected to the PureLink collection tube and the column was disposed of. The purified PCR products were quantified with the NanoDrop 2000 (Thermo Fisher Scientific) and stored at -20 °C. Then, agarose gel electrophoresis was conducted to determine the purified PCR amplicons.

#### 3.2.2.9 DNA sequencing of PCR amplicon and data analysis

The PCR amplicons were sequenced by the Sanger chain termination method with respective primers (both forward and reverse primers) from the DNA sequencing lab of First BASE Laboratories, Malaysia. Obtained DNA sequencing data were visualized from FinchTV v1.4.0 and then assembled, and contigs were created and saved as fasta files using the Unipro UGENE v42 software. DNA sequences to their corresponding peptide sequences were translated by EMBOSS Transeq of the EMBL-EBI web tool (Li *et al.*, 2015). Nucleotide and corresponding polypeptides were identified by BLAST (Basic Local Alignment Search Tool) of NCBI (Johnson *et al.*, 2008) and UniProt database (Zaru & Orchard, 2023). Gene and protein sequences were identified from the GenBank feature of NCBI BLAST alignment. Gene as well as protein features were identified from the GenBank derived window. Obtained protein molecular weights were calculated manually from the protein molecular weight calculator (https://www.bioinformatics.org/sms/prot_mw.html).

## 3.3 Results

### 3.3.1 Hemolytic activity test

Among two hundred ninety-one (291) native *B. thuringiensis* strains studied, 98 were identified as nonhemolytic whereas 193 were found as hemolytic (Figure 3.3.1) on sheep blood agar. All three reference *B. thuringiensis* (kurstaki HD-73, sotto, and japonensis) used in the study were found as hemolytic strains. The statistical presentation of the distribution of hemolytic and nonhemolytic strains of the tested Bt strains are depicted in Figure 3.3.2.

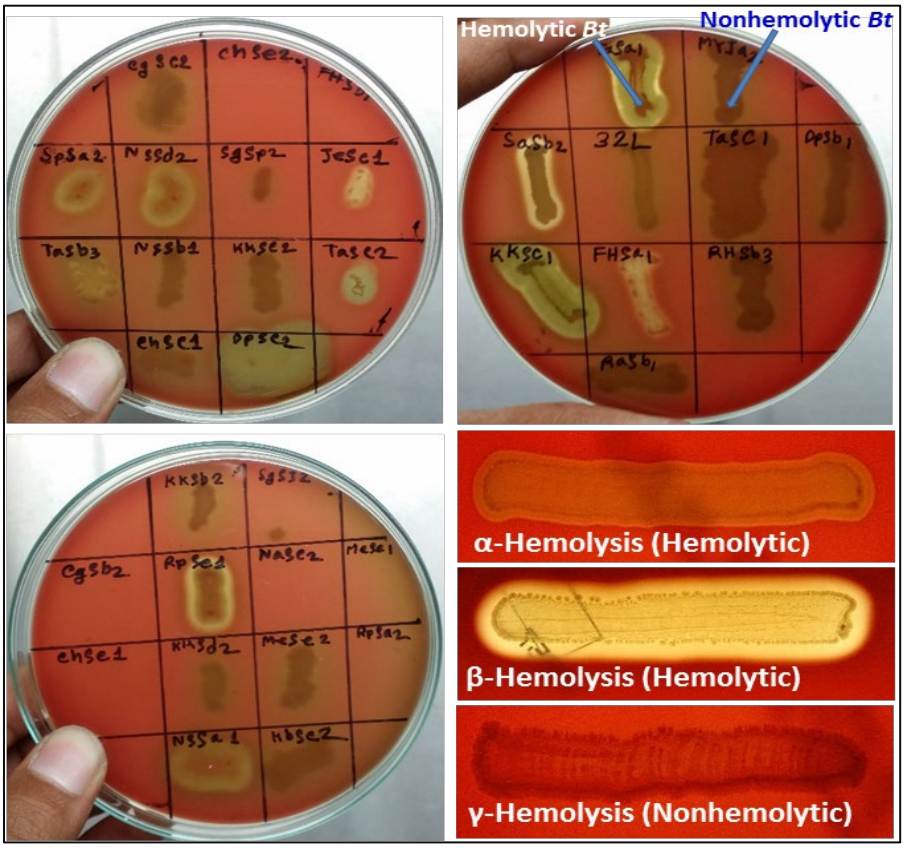



**Figure 3.3.1:** A few typical plates showing the hemolysis (with clear transparent zone) and nonhemolysis (no clear zone) by the Bt strains tested on sheep blood agar.

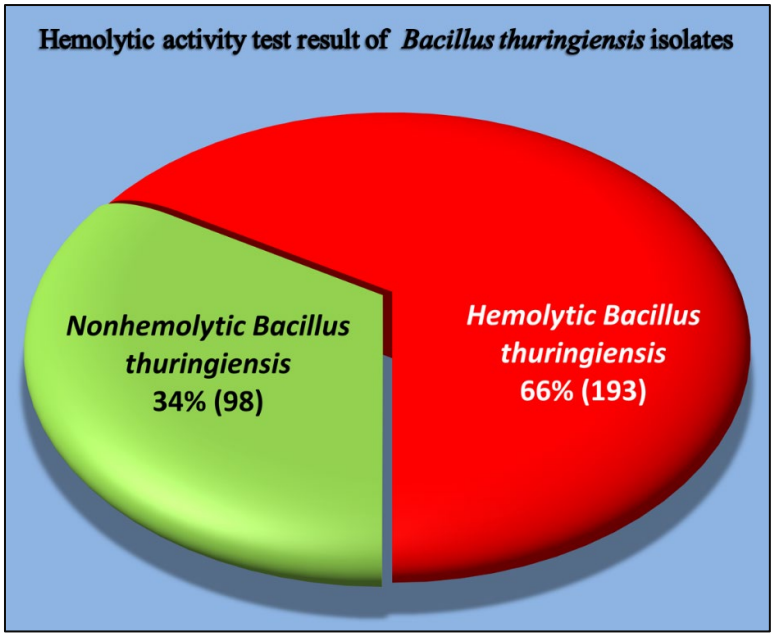



**Figure 3.3.2:** Distribution of hemolytic and nonhemolytic Bt strains amongst the tested indigenous strains.

### 3.3.2 Identification of *ps* genes from nonhemolytic Bt strains by PCR

All of the 98 nonhemolytic Bt strains were screened by polymerase chain reaction (PCR) for detection of the cancer cell-killing parasporal Cry protein synthesizing *ps* genes (*ps1, ps2, ps3, ps4* and *ps5*). There were no amplification products obtained from 73 nonhemolytic Bt strains after the PCR was done with six sets of primer pairs. Specific primers targeting the *ps1, ps3,* and *ps4* genes positively amplified fragments of the predicted sizes from plasmid DNA but also produced some unintended bands Table 3.3.1. No polymerase chain reaction (PCR) amplifications were detected from *ps2* and *ps5*.

Primers specific to the *ps1* gene gave amplification of approximately 1200, 500, 550, 300 & 260 base pair amplicon from the examined Bt strains (Figure 3.3.3 A, B). The 500 base pair PCR products from Bt strains MyIa1, MyIb1, NaSb2, RaSd1, and the 550 bp PCR product from Bt strain FhSb3, which are shown in Figure 3.3.3 (A, B) were considered to be positive. These products were chosen for DNA sequencing since the anticipated size of the ps1 gene amplification product was 542 bp. In addition, despite the PCR amplicons of 1200, 300, and 260 base pairs not being within the expected size, each of them, as exhibited a single band on the agarose gel, was also subjected to DNA sequencing in order to determine their identity.

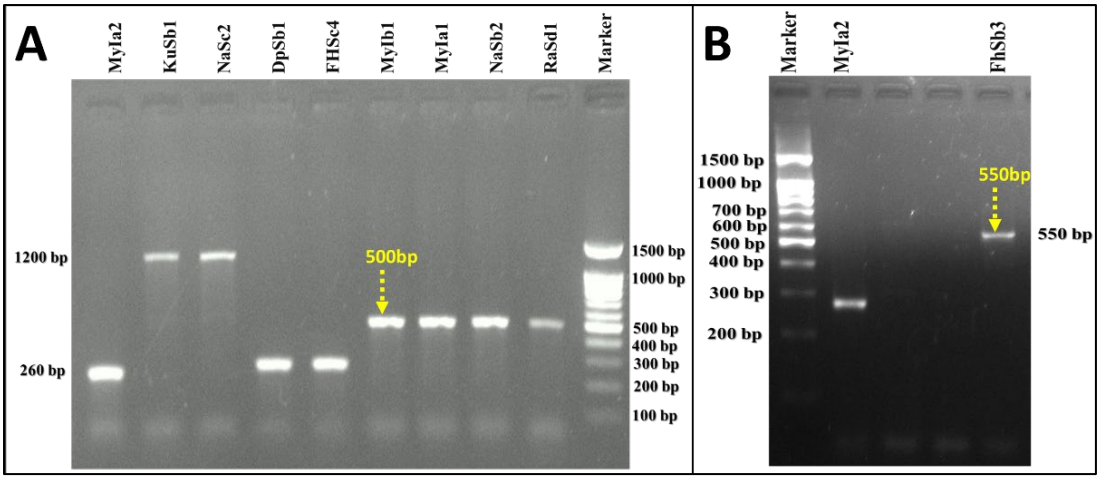



**Figure 3.3.3 :** Representative agarose gel electrophoresis of PCR amplification product of *ps1* gene from nonhemolytic Bt strains (A, B). (Expected size of *ps1*: 542 bp).

The *ps3* (*CryBP*) gene-specific PCR primers successfully generated a PCR amplicon of the anticipated size (499 bp) when plasmid DNA from KkSc1, NaSb2, FhSb3, and SaSe1 strains of Bt were used as template amplifying approximately 500 base pairs amplicons as illustrated in Figure 3.3.4. Consequently, those were subjected to DNA sequencing.

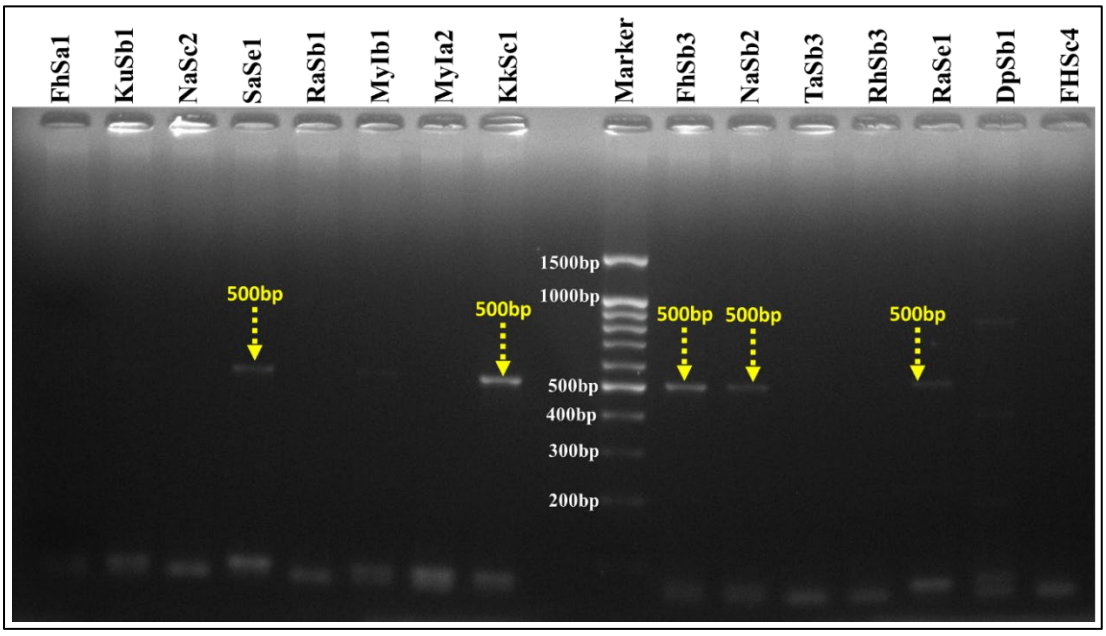



**Figure 3.3.4:** Representative agarose gel electrophoresis of PCR amplification product of *ps*3 gene from nonhemolytic Bt strains. (Expected size of *ps3CryBP*: 499 bp)

When employing plasmid DNA, the *ps4* (*etx-mtx*) gene-specific PCR detection primers amplified a 700 bp PCR product from five Bt strains, namely FhSc4, KbSa1, KkSc1, MyIa2, and NaSb1 (Figure 3.3.5 A, B). Despite the anticipated length of the

PCR amplicon is 388 bp, the PCR products were selected for sequencing to ascertain their source, as distinct bands were detected using *ps4 (etx-mtx)* gene-specific primers.

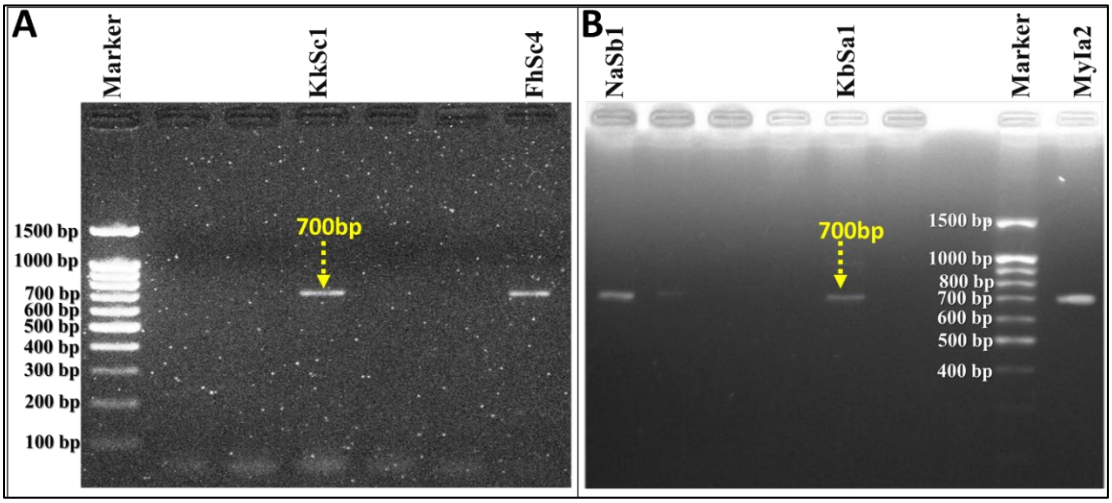



**Figure 3.3.5:** PCR amplicons of native nonhemolytic *B. thuringiensis* strains on representative agarose gel electrophoresis with parasporin-4 (*etx-mtx*) detection primer. (Expected PCR amplicon: 388 bp).

The size of the amplified PCR products using plasmid DNA extracted from the listed Bt strains, which were subjected to sequencing, are shown in Table 3.3.1.

**Table 3.3.1:** Summary of the PCR amplicon obtained from indigenous nonhemolytic Bt strains

| Serial | Bt strains | *ps1* (*δ-endotoxin N*) amplicon size (bp) | *ps3* (*CryBP1*) amplicon size (bp) | *ps4* (*etx-mtx*) amplicon size (bp) |
|---|---|---|---|---|
| 1. | DpSb1 | 300 | | |
| 2. | FhSb3 | 550 | 500 | |
| 3. | FhSc4 | 300 | | 700 |
| 4. | KbSa1 | | | 700 |
| 5. | KkSc1 | | 500 | 700 |
| 6. | KuSb1 | 1200 | | |
| 7. | MyIa1 | 500 | | |
| 8. | MyIa2 | 260 | | 700 |
| 9. | MyIb1 | 500 | | |
| 10. | NaSb1 | | | 700 |
| 11. | NaSb2 | 500 | 500 | |
| 12. | NaSc1 | 1200 | | |
| 13. | RaSd1 | 500 | | |
| 14. | RaSe1 | | 500 | |
| 15. | SaSd1 | | | |
| 16. | SaSe1 | | 500 | |

(*ps1* desired size: 542 bp; *ps3* desired size: 499 bp; *ps4* desired size: 388 bp)

### 3.3.3 Identification of *ps* genes from hemolytic Bt strains by PCR

The plasmid and genomic DNA from all of the 23 indigenous Bt strains, which are hemolytic but non-insecticidal, as well as 3 reference Bt strains, were analyzed using six sets of primers specifically designed to identify parasporin genes (Table 3.2.3). The Bt strains underwent screening to see whether they had the *ps* genes (*ps1, ps2, ps3, ps4*, and *ps5*). The *ps1, ps3,* and *ps4* genes were selectively amplified from plasmid DNA using specific primers (Table 3.3.2). Both expected and unexpected bands were amplified from *ps1, ps3,* and *ps4*, but no PCR amplifications were observed from *ps2* and *ps5*.

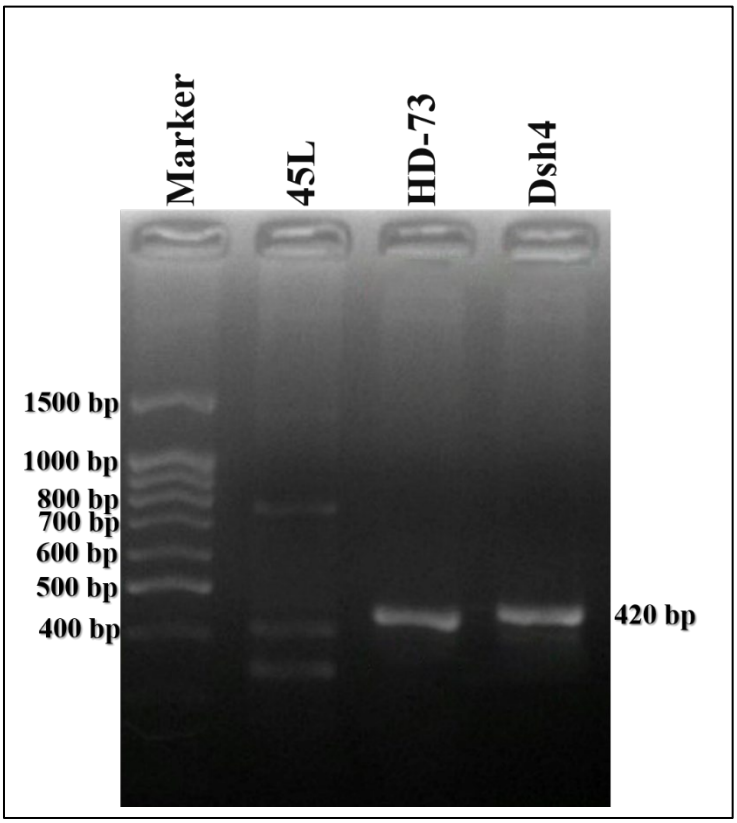



**Figure 3.3.6:** Agarose gel electrophoresis of the amplified *ps1* gene specific primer from hemolytic Bt strains. (Expected PCR amplicon size of *ps1*: 542 bp)

Primers specific to *ps1* gene gave amplification of approximately 420 bp amplicons from two among all the 23 indigenous and 3 references hemolytic Bt strains (Figure 3.3.6). The expected size of the *ps1* gene amplification product was 542 bp. Although the 420 base pair PCR products obtained from the Bt strain DSh4 and the reference HD-73, do not fall within the expected range, to ascertain their identity and origin of purified PCR amplified products from these reactions were selected for sequencing.

Besides, multiple bands were seen in the PCR using strain 45L as a template, which seems nonspecific amplification; hence no sequencing was performed for these PCR amplicons.

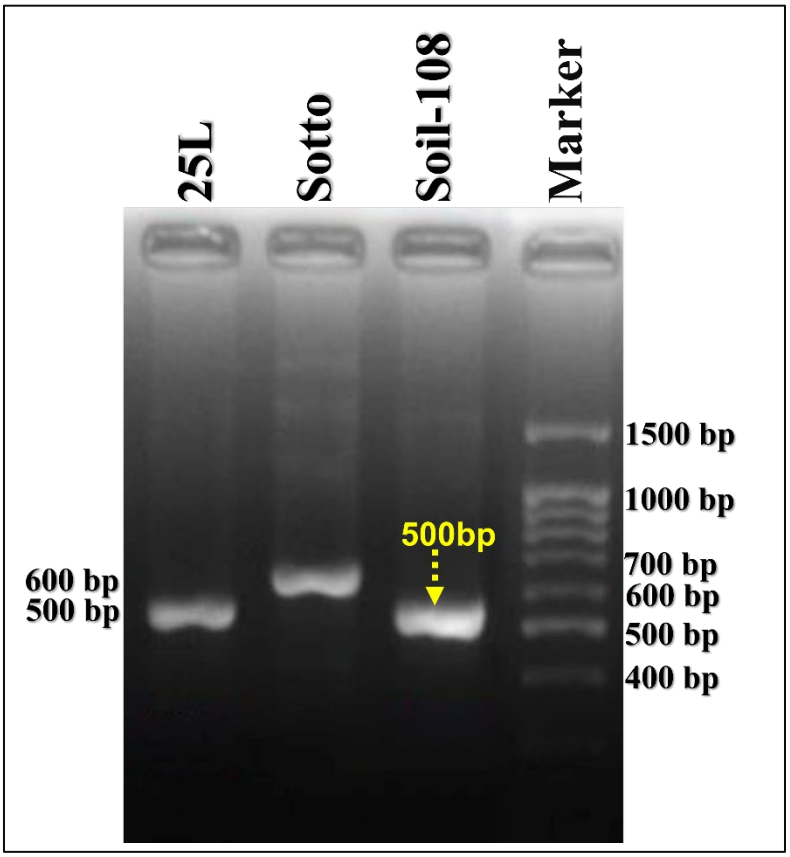



**Figure 3.3.7:** Agarose gel electrophoresis of the amplified *ps3* (*CryBP*) gene specific primer from hemolytic Bt strains. (Desired PCR amplicon size *CryBP* of *ps3*: 499 bp)

All of the 23 indigenous and 3 reference strains were checked for the detection of parasporin-3. The PCR primers designed for the *ps3* (*CryBP*) gene effectively produced a PCR amplicon of the expected size using plasmid DNA from two hemolytic non-insecticidal indigenous Bt strains. PCR amplifications of approximately 500 base pairs were detected from agarose gel electrophoresis of 25L and Soil-108, as shown in Figure 3.3.7. The expected size of the *ps3* gene amplification product is 499 bp. Since the PCR amplicons from both strains matched the appropriate size, the PCR result from both strains was selected for DNA sequencing. In addition, a 600 bp PCR amplicon from the reference Bt strain sotto was also sent for sequencing despite the amplification of an unexpected-sized product.

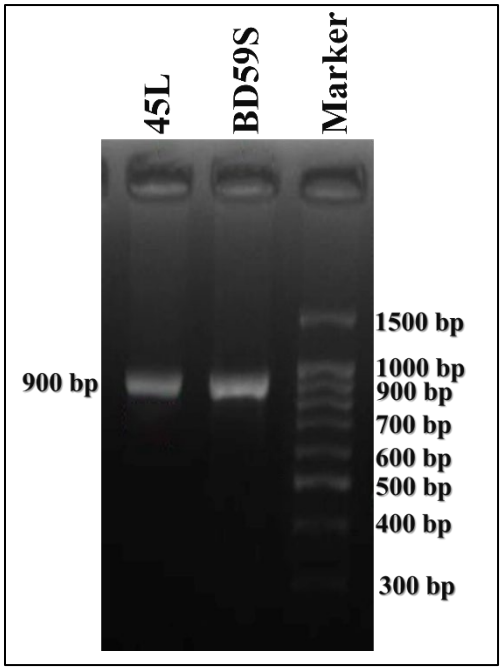



**Figure 3.3.8:** Agarose gel electrophoresis of the amplified *ps4* (*etx-mtx*) gene specific primer from hemolytic Bt strains. (Desired PCR amplicon size of *etx-mtx* of *ps4*: 388 bp)

The presence of parasporin-4 was assessed in all 23 indigenous and 3 reference hemolytic Bt strains. When using plasmid DNA, the *ps4* (*etx-mtx*) gene-specific PCR detection primers produced a PCR amplicon with a size of around 900 base pairs from two Bt strains, namely 45L and BD59S (Figure 3.3.8). Nevertheless, the desired or appropriate size of the PCR band should be 388 base pairs. The PCR results that were obtained from both Bt strains were submitted for DNA sequencing even though the amplification of an unexpectedly large product occurred.

**Table 3.3.2:** Summary of the PCR amplicon result of indigenous and reference hemolytic *Bt* strains

| **Bt strains** | ***ps1* (*δ-endotoxin N*) amplicon size (bp)** | ***ps3* (*CryBP1*) amplicon size (bp)** | ***ps4* (*etx-mtx*) amplicon size (bp)** |
|---|---|---|---|
| 25L | | 500 | |
| 45L | Multiple | | 900 |
| BD59S | | | 900 |
| Dsh4 | 420 | | |
| HD-73 | 420 | | |
| Soil-108 | | 500 | |
| Sotto | | 600 | |

(*ps1* desired size: 542 bp; *ps3* desired size: 499 bp; *ps4* desired size: 388 bp)

### 3.3.4 DNA sequenced data analysis of PCR amplicons from nonhemolytic Bt strains

The available parasporin sequences in the databases of GenBank were limited when this study was initiated. In fact, the deposited parasporin sequences did not increase either in a tangible fashion which would ensure the probability of achieving a detectable amplification of the putative sequence. The illustrated results in this investigation showed that none of the six designed primer sets could successfully amplify potential parasporin sequences. Out of the 98 strains used in this study, only 16 strains possessed putative parasporin gene fragments as compared with the preset features. The PCR amplified sequence lengths varied from as small as 260 bp up to 1200 bp as resolved using conventional agarose gel electrophoresis. DNA sequencing was performed for 18 PCR amplicons of indigenous nonhemolytic *B. thuringiensis* strains, of them 10 from *ps1,* 4 from *ps3*, and 4 from *ps4* primer-specific amplicon.

Obtained DNA sequencing data were observed from FinchTV v1.4.0 and were found to be good in quality. Then, contig was created using the Unipro UGENE v42 software and saved as fasta file. A total of 18 contigs were found from 18 PCR amplicon sequencing data. The resulting fasta contigs were identified from nucleotide BLAST in the NCBI database. Some of the sequence files were found to match part of genomes or part of plasmids of *B. cereus* group or *B. thuringiensis* bacteria (Figure 3.3.9 A, B). So, the GenBank feature of the NCBI nucleotide BLAST alignment was checked, and from there, gene name, as well as protein name, ID and sequence, was obtained. Unfortunately, the BLAST results were not in agreement with any of the parasporins or cry genes present in the database. A brief description of the identified proteins with their molecular weight, source, and accession numbers obtained from the DNA sequencing data analysis is shown in Table 3.3.3.

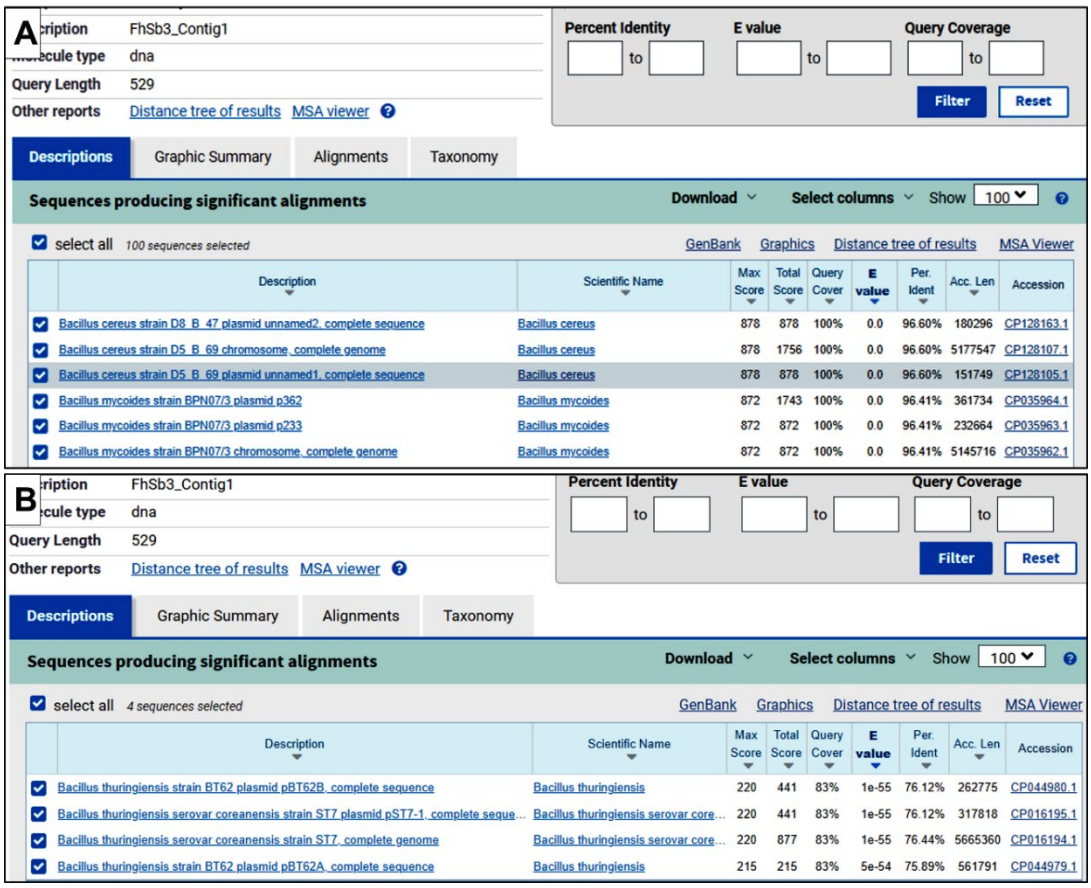



**Figure 3.3.9:** Identification of *ps1* specific PCR amplified DNA sequencing data nonhemolytic Bt strains from NCBI nucleotide BLAST. (A) BLAST result with *Bacillus cereus* group, (B) BLAST result with *B. thuringiensis* database.

Among DNA sequencing data of 10 PCR amplicons of *ps1* gene-specific primer, the agarose gel observed 550 bp PCR amplicon from Bt FhSb3 was identified as a 529 nucleotide long partial gene sequence of a *ltrA* gene, which is responsible for synthesizing a mobile ribozyme, called group II intron reverse transcriptase/maturase. These PCR products were chosen for DNA sequencing since the anticipated size of the *ps1* gene amplification product was 542 bp, but no *ps* genes, as well as other cry genes were identified from it (Figure 3.3.9 A, B). Besides, from NCBI nucleotide BLAST of the 260 bp PCR sequencing data, an 11.45 kDa multispecies hypothetical protein synthesizing gene was identified. In contrast, the 300 bp was identified as a 33.71 kDa Drug/metabolite transporter (DMT) family protein synthesizing gene. Besides, the 500 bp PCR sequencing data was identified as a 79.85 kDa Zinc-ribbon domain-containing protein synthesizing gene, and the 1200 bp was identified as a Lamin tail, metallophosphatase and ribonuclease domain containing 209.11 kDa protein

synthesizing gene of *B. thuringiensis*. DNA sequencing results of *ps1* primer-specific amplified data are shown in Table 3.3.3.

The agarose gel observed 500 bp PCR amplicons of *ps3* gene-specific primer from four Bt strains, KkSc1, NaSb2, FhSb3, and SaSe1, was identified as a 515 nucleotide long partial gene sequence of a rare *CheR* gene rather than any *ps* or *cry* gene. When BLAST was performed, only a single match was found with *Bacillus weihaiensis* strain Alg07. Thus, this nucleotide sequence was then translated into protein and again, BLAST was performed on NCBI protein BLAST (Figure 3.3.10). This gene is responsible for the synthesis of protein-glutamate O-methyltransferase CheR, which is a chemotaxis protein of *B. thuringiensis* and multi-species bacteria (Table 3.3.3). These PCR products were chosen for DNA sequencing since the anticipated size of the *ps3* gene amplification product was 499 bp. Still, no *ps* genes, as well as other cry genes, were identified from the 515-nucleotide long DNA sequencing data of PCR products from 4 different Bt strains.

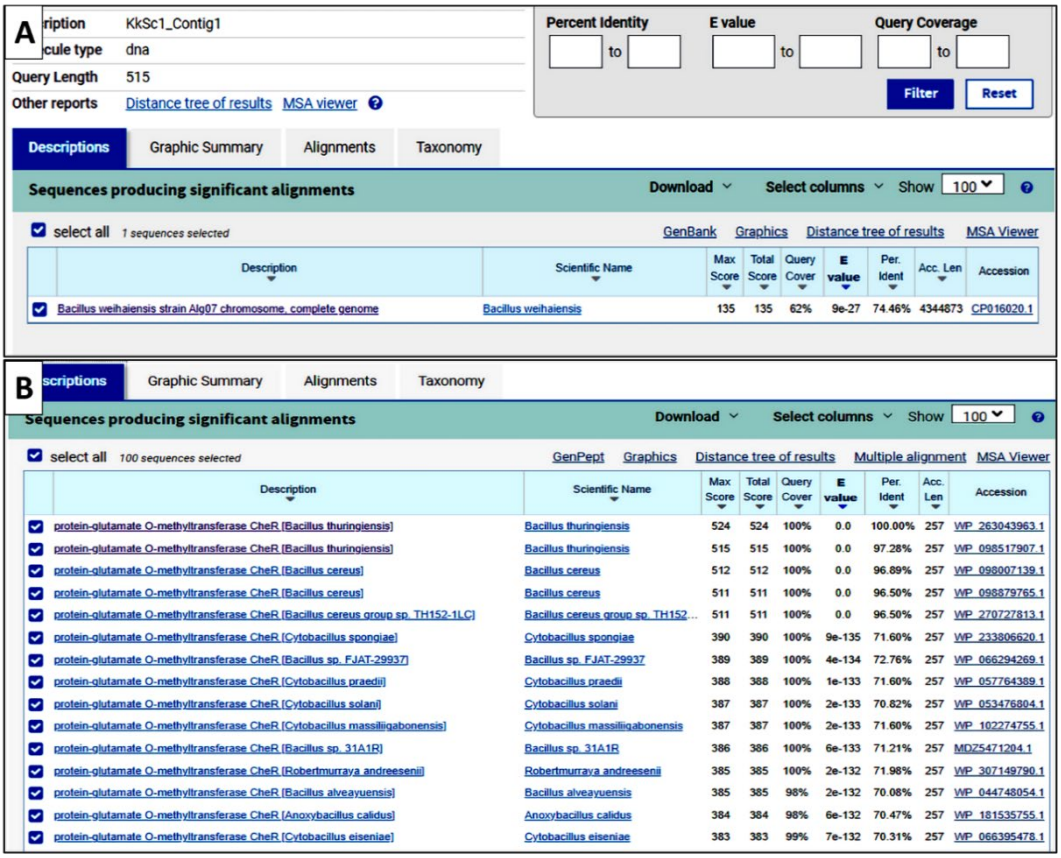



**Figure 3.3.10:** Identification of *ps3* specific PCR amplified DNA sequencing data nonhemolytic Bt strains from NCBI BLAST, A: nucleotide BLAST, B: protein BLAST.

The agarose gel observed 700 bp PCR amplicons of *ps4* gene-specific primer from four Bt strains KkSc1, FhSc4, MyIa2, and KbSa1 were identified from NCBI nucleotide BLAST as a 607 nucleotide long partial gene sequence of a Tn3 family transposase gene rather than any *ps* or *cry* gene (Table 3.3.3). When BLAST was performed, it was found that this gene is located in the plasmid of *B. thuringiensis* and *Enterococcus* bacteria (Figure 3.3.11).

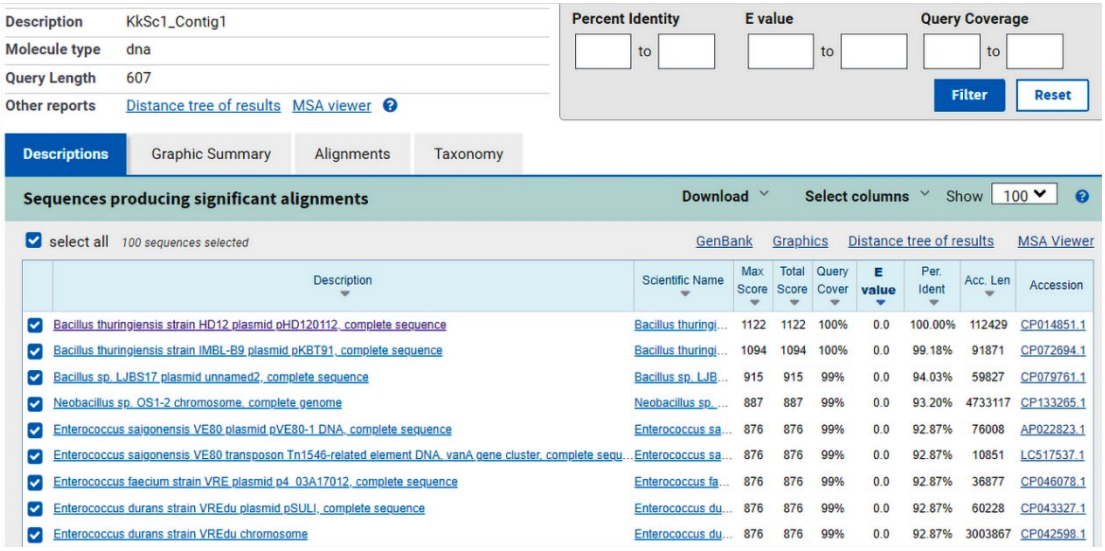



**Figure 3.3.11:** Identification of *ps4* specific PCR amplified DNA sequencing data of nonhemolytic Bt strains from NCBI nucleotide BLAST.

**Table 3.3.3:** Summary of the DNA sequencing result of parasporin gene-specific PCR amplicons of indigenous nonhemolytic Bt strains

| PCR amplicon size & Bt strains | Identified protein molecules of DNA sequencing result from NCBI BLAST |
|---|---|
| ***ps1*-260 bp MyIa2** | 11.45_kDa Multispecies Hypothetical protein (MBE5107852.1) of *B. thuringiensis* |
| ***ps1*-300 bp DpSb1** | 33.71_kDa Drug/metabolite transporter (DMT) family protein (WP_088119307.1) of *B. thuringiensis* |
| ***ps1*-300 bp FhSc4** | 33.71_kDa Drug/metabolite transporter (DMT) family protein (WP_088119307.1) of *B. thuringiensis* |
| ***ps1*-500 bp RaSd1** | 79.85_kDa Zinc-ribbon domain-containing protein (WP_088021476.1), 79.83 kDa hypothetical protein (WP_002204126.1) of *B. thuringiensis* |
| ***ps1*-500 bp MyIa1** | 79.85_kDa Zinc-ribbon domain-containing protein (WP_088021476.1), 79.83 kDa hypothetical protein (WP_002204126.1) of *B. thuringiensis* |

| PCR amplicon size & Bt strains | Identified protein molecules of DNA sequencing result from NCBI BLAST |
|---|---|
| ***ps1*-500 bp MyIb1** | 79.85_kDa Zinc-ribbon domain-containing protein (WP_088021476.1), 79.83 kDa hypothetical protein (WP_002204126.1) of *B. thuringiensis* |
| ***ps1*-500 bp NaSb2** | 79.85_kDa Zinc-ribbon domain-containing protein (WP_088021476.1), 79.83 kDa hypothetical protein (WP_002204126.1) of *B. thuringiensis* |
| ***ps1*-550 bp FhSb3** | 65.1_kDa Group II intron reverse transcriptase/maturase (MBE5109091.1) of *B. thuringiensis* |
| ***ps1*-1200 bp NaSc2** | 209.11_kDa Lamin tail domain-containing protein also have metallophosphatase and ribonuclease E domain (WP_233319428.1) of *B. thuringiensis* |
| ***ps1*-1200 bp KuSb1** | 209.11_kDa Lamin tail domain-containing protein also have metallophosphatase and ribonuclease E domain (WP_233319428.1) of *B. thuringiensis* |
| ***ps3*-500 bp FhSb3** | 30.58_kDa Protein-glutamate O-methyltransferase CheR (WP_263043963.1), Chemotaxis protein, nucleoside-diphosphate kinase of *B. thuringiensis* |
| ***ps3*-500 bp KkSc1** | 30.58_kDa Protein-glutamate O-methyltransferase CheR (WP_263043963.1), Chemotaxis protein, nucleoside-diphosphate kinase of *B. thuringiensis* |
| ***ps3*-500 bp NaSb2** | 30.58_kDa Protein-glutamate O-methyltransferase CheR (WP_263043963.1), Chemotaxis protein, nucleoside-diphosphate kinase of *B. thuringiensis* |
| ***ps3*-500 bp SaSe1** | 30.58_kDa Protein-glutamate O-methyltransferase CheR (WP_263043963.1), Chemotaxis protein, nucleoside-diphosphate kinase of *B. thuringiensis* |
| ***ps4*-700 bp KkSc1** | 114.57_kDa Tn3 family transposase (AMR88290.1) of *B. thuringiensis* |
| ***ps4*-700 bp FhSc4** | 114.37_kDa Tn3 family transposase (WP_048557829.1) of *B. thuringiensis* |
| ***ps4*-700 bp MyIa2** | 114.37_kDa Tn3 family transposase (WP_048557829.1) of *B. thuringiensis* |
| ***ps4*-700 bp KbSa1** | 114.37_kDa Tn3 family transposase (WP_048557829.1) of *B. thuringiensis* |

### 3.3.5 DNA sequencing data analysis result of PCR amplicon of non-insecticidal hemolytic Bt strains

DNA sequencing was performed for 7 PCR amplicons of indigenous non-insecticidal but hemolytic *B. thuringiensis* strains, of them 2 from *ps1,* 3 from *ps3*, and 2 from *ps4*

primer-specific amplicon. A total of 7 contigs were found from 7 PCR amplicon sequencing data. The resulting fasta contigs were identified from nucleotide BLAST in the NCBI database. Like DNA sequence data of nonhemolytic Bt, no gene name was directly predicted from the BLAST result of hemolytic Bt (Figure 3.3.12). So, the gene and protein sequence were obtained from the GenBank feature of the NCBI nucleotide BLAST alignment. Unfortunately, no expected parasporin and any other cry gene were identified from the BLAST result of the hemolytic Bt source. A summarized description of the identified proteins with their molecular weight, source, and accession numbers obtained from the DNA sequencing data analysis is shown in Table 3.3.4. The agarose gel observed 500 bp PCR amplicon of *ps3* gene-specific primer from 25L, and Soil-108 was identified as a 519 nucleotide long partial sequence of a gene that encoded polysaccharide deacetylase family protein in *B. thuringiensis* strains. Besides, a gene that encoded DMT family transporter protein was identified from DNA sequencing data of the *ps4*-specific PCR amplicon.

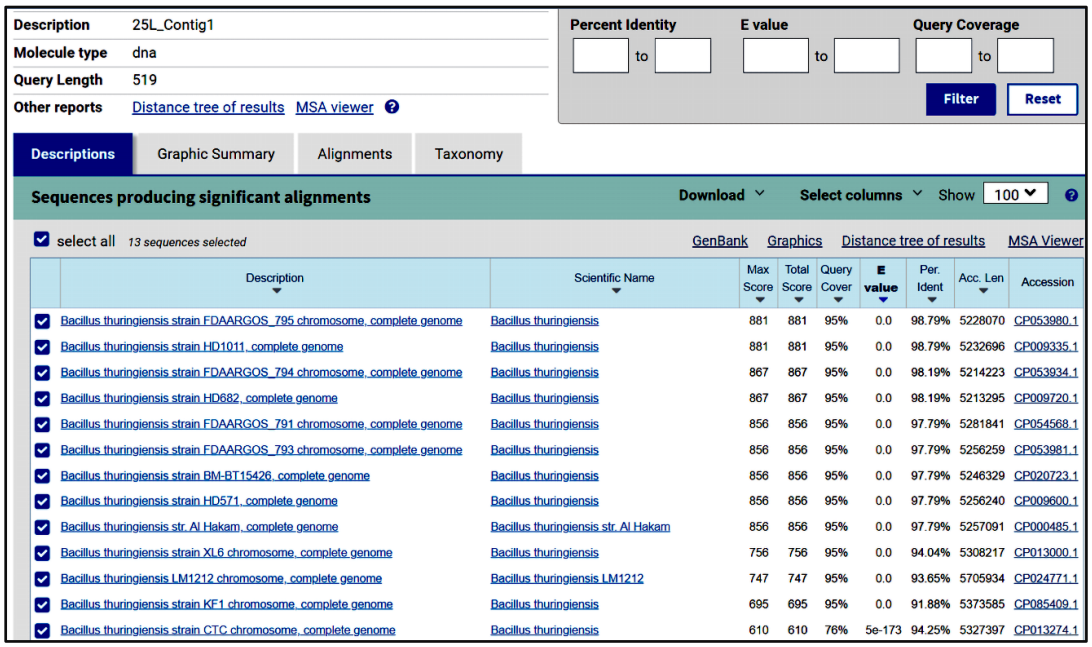



**Figure 3.3.12:** Identification of *ps3* specific PCR amplified DNA sequencing data of hemolytic Bt strains from NCBI nucleotide BLAST.

**Table 3.3.4:** Summary of the DNA sequencing result of parasporin gene-specific PCR amplicons of non-insecticidal hemolytic Bt strains

| PCR amplicon size & Bt strains | Identified protein molecules of DNA sequencing result from NCBI BLAST |
|---|---|
| ***ps1*-420 bp Dsh4** | 7.87_kDa hypothetical protein, (WOE47724.1) and 9.15_kDa DUF3961 domain-containing protein (WOE47725.1) of *B. thuringiensis*. Primers primed in the upstream and downstream sequence of these two protein coding gene |
| ***ps1*-420 bp HD-73** | 7.87_kDa hypothetical protein, (WOE47724.1) and 9.15_kDa DUF3961 domain-containing protein (WOE47725.1) of *B. thuringiensis*. Primers primed in the upstream and downstream sequence of these two protein coding gene |
| ***ps3*-500 bp 25L** | Bt 42.12_kDa polysaccharide deacetylase family protein (QKH25340.1) of *B. thuringiensis* |
| ***ps3*-500 bp Soil-108** | Bt 42.12_kDa polysaccharide deacetylase family protein (QKH25340.1) of *B. thuringiensis* |
| ***ps3*-600 bp Sotto** | Multidrug transporter (ALL22810.1), EamA family transporter (ATI51685.1) of *B. thuringiensis* |
| ***ps4*-900 bp 45L** | 33.7_kDa DMT family transporter (MDC2944555.1) of *B. thuringiensis* |
| ***ps4*-900 bp BD59S** | 33.7_kDa DMT family transporter (QDQ05016.1) of *B. thuringiensis* |

In this study, we explored the presence of parasporin genes, which are known to be a rare gene, in *B. thuringiensis* isolated from our native soil. Although none of the sequences showed similarity with the known parasporin genes of the existing database, we did find 11 different types of genes on the plasmids. The majority of the genes amplified by parasporin-specific primers encode genes important in regulating cell growth and maintenance, for example, bacterial membrane-related transporter protein, bacterial transposable element, and enzymes engaged in the manipulation of DNA and RNA. In summary, while it is scientifically logical to hypothesize that proteins encoded by genes from *B. thuringiensis* could play roles in killing human cancer cells, extensive research and validation are required to confirm this potential and understand the underlying mechanisms.

# CHAPTER 4:

# Isolation of parasporal anticancer proteins from indigenous nonhemolytic *Bacillus thuringiensis* strains

# CHAPTER 4: Isolation of parasporal anticancer proteins from indigenous nonhemolytic *Bacillus thuringiensis* strains

## 4.1 Introduction

The parasporal Cry proteins of insecticidal *B. thuringiensis* are promising agents for the microbiological control of multiple orders of insects (Cannon, 1996; Lambert & Peferoen, 1992). Cry proteins are also found in the non-insecticidal *B. thuringiensis* strains (Hastowo *et al.,* 1992; Meadows *et al.,* 1992; Ohba, 1996). Solubilized non-insecticidal Cry proteins of some *B. thuringiensis* strains, when digested with proteinase K or trypsin have shown human cancer cell-specific anticancer activity (Gonzalez *et al.,* 2011; Jung *et al.,* 2007; Lenina *et al.*, 2014; Mizuki *et al.*, 1999; Nagamatsu *et al.,* 2010; Ohba *et al.,* 2009; Poornima *et al.,* 2010; Uemori *et al.*, 2007; Wong., 2010; Yasutake *et al.*, 2006). These types of Cry proteins are called parasporins. These parasporins are divided into three categories: three-domain type *α-PFTs* (pore-forming toxins), Cyt toxin type β-PFTs, and aerolysin type β-PFTs (Xu *et al.* 2014). Moreover, based on amino acid sequence, molecular weight, structure, and target specificity, these proteins have also been categorized into 6 (six) classes, parasporin (PS1 to PS6). Each class has potential and selective cytotoxic activity against different human carcinoma cells and has been isolated mainly from Japan (Uemori *et al.*, 2008), Vietnam (Yasutake *et al.*, 2006), and Canada (Jung *et al.*, 2007). There are also reports of the occurrence of the parasporins from soils in India (Lenina *et al.*, 2014; Poornima *et al.*, 2010) and Caribbean Island (Gonzalez *et al.*, 2011).

As parasporins of *B. thuringiensis* have selective anticancer activities against cancer cell lines, they could play an important role in targeted peptide therapy. Parasporin-1 was isolated from nonhemolytic *B. thuringiensis* strains (Ichikawa *et al.*, 2008). Despite the findings from the previous chapter, where amplified gene fragments through PCR showed no similarity with the existing database of parasporin genes or other Cry protein-encoding genes, the study was directed to explore the novel anticancer parasporal inclusion proteins of indigenous nonhemolytic *B. thuringiensis* strains.

## 4.2 Materials and Methods

### 4.2.1 Materials

#### 4.2.1.1 Chemicals

Acrylamide ($C_3H_5NO$), agar ($C_{14}H_{24}O_9$), ammonium persulfate (APS) {$(NH_4)_2S_2O_8$}, n,n′-methylenebisacrylamide $CH_2[NHC(O)CH=CH_2]_2$, bromophenol blue ($C_{19}H_{10}Br_4O_5S$), coomassie blue G-250 ($C_{47}H_{48}N_3NaO_7S_2$), $dH_2O$, DTT (dithiothreitol) ($C_4H_{10}O_2S_2$), glacial acetic acid ($CH_3COOH$), glycine ($C_2H_5NO_2$), phenylmethylsulfonyl fluoride (PMSF) ($C_7H_7FO_2S$), $Na_2EDTA$ ($C_{10}H_{14}N_2O_8.2Na.2H_2O$), ethanol ($CH_3CH_2OH$), n-hexane ($C_6H_{14}$), glycerol, HCl, methanol ($CH_3OH$), $MnCl_2$, KCl, NaCl, $K_2HPO_4$, $Na_2HPO4$, $KH_2PO4$, $Na_2CO_3$, phosphoric acid ($H_3PO_4$), SDS {$CH_3(CH_2)_{11}OSO_3Na$}, tris-base ($C_4H_{11}NO_3$). The chemicals used in this study were of molecular grade, and their lists and company names are mentioned in Appendix **III**.

#### 4.2.1.2 List of Equipment

Autoclave machine, biosafety cabinet, centrifuge machine, electronic balance, microbiological incubator, magnetic stirrer, micropipettes, orbital shaker incubator, pH meter, phase-contrast microscope, SDS gel electrophoresis power supply, -20 $^oC$ fridge, 4 $^oC$ refrigerator, sonication machine, spectrophotometer, thermo stated shaking water bath, vortex mixture, etc., apparatuses used in this study are mentioned in the respective methods section, and their company and models were mentioned in **Appendix IV**.

#### 4.2.1.3 Media

Luria Bertani (LB) agar and broth, and T3 broth, were used in this study, and their compositions are mentioned in **Appendix I**.

#### 4.2.1.4 Reagents and Solutions

Bacto-tryptone, bacto-tryptose, bromophenol blue (1%), yeast extract, proteinase K, phosphate-buffer, phosphate-buffered saline (PBS), normal saline, Tris-HCl, EDTA, parasporal protein solubilizing buffer (pH 11.0), Bradford reagent, proteinase K dissolving solution, 0.1 M PMSF, 30% acrylamide-bisacrylamide, 10% APS, SDS upper gel buffer (pH 6.8), SDS lower gel buffer (pH 8.8), SDS electrophoresis buffer, SDS staining and destaining solution, protein sample preparation buffer, etc., reagents

and solutions used in these experimental procedures were mentioned in the respective methods section. Their composition and preparation methods are mentioned in **Appendix II.**

#### 4.2.1.5 Bacterial strains

This research was carried out by screening 98 indigenous nonhemolytic *B. thuringiensis* strains (Table 3.2.1), previously isolated from different locations and ecosystems in Bangladesh.

### 4.2.2 Methods

#### 4.2.2.1 Maintenance of Culture

Bt strains were maintained in LB agar by subculturing and stored at 4 $^{o}$C for further work.

#### 4.2.2.2 Parasporal crystal protein preparation

To get pure parasporal crystalline proteinaceous inclusions (without contamination of any membrane or vegetative cell protein) and to find out the Cry synthesizer, completely sporulated *B. thuringiensis* culture was required, an approach to develop the best cultural condition. For this reason, a pure and single reference bacterial inoculum (HD73) was incubated in 50 mL of T3, LB, and LB plus phosphate buffer (pH 7), these 3 different broths separately at 30 °C in an orbital shaking incubator (Excella E25 Incubator Shaker series, New Brunswick Scientific) at 150 rpm. Cells were grown for about 3–7 days until complete synthesis of the spores and the parasporal Cry. After the 3rd, 5th, and 7th days, 100 µL of culture was collected from each T3, LB, and LB plus phosphate buffer broth and checked under a phase-contrast microscope (Primo Star, ZEISS). After the 7th day, the culture was then kept in the refrigerator at 4 $^{o}$C for 16–18 hours and finally observed under a phase-contrast microscope. In this step, completely sporulated *B. thuringiensis* culture was found only in T3 broth, and this condition was considered an ideal cultural condition for further experiments.

#### 4.2.2.3 Spore crystal separation using the n-hexane biphasic separation method

The Crystal proteins were separated from spore mixtures using n-hexane, a previously described method (Mounsef *et al.*, 2014). Briefly, the procedure was, a spore crystal

mixture from completely sporulated Bt bacteria of the T3 broth was collected in separate 50 mL falcon tubes for each strain by centrifugation at 6000 × g at 4 °C for 10 minutes. The obtained pellet was washed twice by suspending it in sterilized ice-cold PBS. Then 10% of n-hexane (purity 95%) was added as an organic solvent (1 mL per 100 mL of aqueous suspension) to minimize the risk of altering the protein crystals. The suspension was sonicated at 100 watts for 10 minutes with a 10 second pulse (Omni-Ruptor 4000 Ultrasonic Homogenizer, OMNI International, USA) to disperse the bacterial spore-crystal clumping and then centrifuged at 5800 × g for 10 minutes. After removing the n-hexane, the pellet was resuspended in 1X sterilized PBS solution, and the process was repeated three or four times. At last, icy distilled water was used to wash the pellet twice.

#### 4.2.2.4 Alkali solubilization of parasporal Cry proteins

The n-hexane purified Crystal proteins were mixed well in 600 µL of freshly prepared 56 mM $Na_2CO_3$ (pH 11.0) solution. Then 12 µL of 1 M dithiothreitol (DTT) was added to each solution and incubated for 2 hours at 37 °C in a water bath (Water Bath 1083, GFL). After 2 hours of incubation, insoluble materials and spores were pelleted by centrifugation at 15800 × g for 2 minutes. The supernatants from each sample were then passed through a 0.22 µm membrane filter. 250 µL of the filtrate from each sample was transferred to different sterile 1.5 mL microcentrifuge tubes, and the pH was adjusted to 8.0 with 1 M Tris-HCl (pH 5.32) or phosphate buffer (pH 5.0). To confirm the presence of the parasporal proteins, an SDS-PAGE analysis was performed. As cry proteins and DTT are light-sensitive (Pozsgay *et al.,* 1987), all of the falcon and microcentrifuge tubes were wrapped with aluminum foil.

#### 4.2.2.5 Bradford protein assay

Solubilized parasporal protein concentration was determined by the previously described Bradford method (Bradford, 1976). The Bradford assay relies on the change in light absorption characteristics of the dye Coomassie Brilliant Blue G-250 caused by the binding of the protein. The dye has the greatest absorption at a wavelength of 465 nm when it is dissolved in an acidic solution (85% phosphoric acid). The maximum absorption of the dye is changed to occur at 595 nm when protein is added and the light absorbance at 595 nm also rises linearly with protein concentration. This increase in absorbance can be measured with a spectrophotometer (Genesys 5, Thermospectronic).

A standard curve of protein concentration vs absorbance at 595 nm was generated by serially diluting a stock of BSA protein and plotting the results (Figure 4.2.1).

The equation **y = mx** represents the relationship between absorbance (y) at 595 nm and protein concentration (x) in mg/mL. The slope of the standard curve (m) determines the relationship between absorbance and protein concentration.

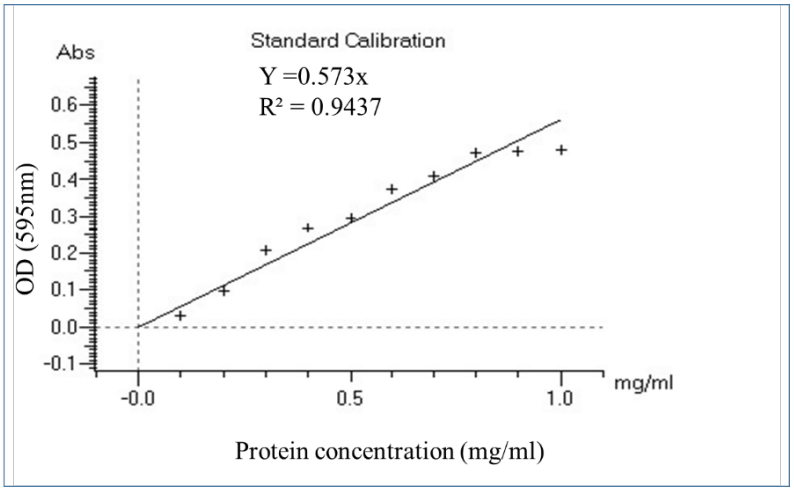



**Figure 4.2.1:** Standard curve for determination of protein concentration by Bradford method

1 mL of Bradford reagent was taken to a 1.5 mL microfuge tube where 20 µL of distilled water was added to use as a control. In the test reaction 20 µL parasporal protein solution instead of water. The tubes were allowed to stand at room temperature for five minutes. The microfuge tube was agitated by inverting it many times to achieve thorough mixing. Subsequently, the measurement of absorbance was conducted against the control sample, and the protein concentration of the test sample was quantified using the standard curve (Chapter 4, and Chapter 5).

#### 4.2.2.6 Proteolytic digestion, activation, and sterilization of parasporal Cry proteins

Each of the solubilized cry protein samples was digested with the proteolytic enzyme, proteinase K (final concentration of 185 µg/mL) by incubating for about 1 hour at 37°C in a water bath (Water Bath 1083, GFL). After that, phenylmethylsulfonyl fluoride (PMSF) was added (final concentration: 1 mM) to each cry protein sample to stop the proteolytic processing. Then, each sample was sterilized with a 0.22 µm membrane filter. All of the microcentrifuge tubes containing parasporal proteins were wrapped

with aluminium foil to protect them from light.

#### 4.2.2.7 SDS-PAGE analysis

To confirm the presence of the parasporal proteins and to determine their molecular weight, SDS-PAGE analysis was performed by using a 4% stacking and 10% separating gel environment.

#### 4.2.2.8 Preparation of SDS polyacrylamide gel

The polyacrylamide gel casting platform included clean, dry glass plates joined with spacers (the Bio-Rad instrument). Distilled water, lower gel buffer, 30% acrylamide and bis-acrylamide solution (29:1), 10% APS, and TEMED were mixed gently to create a 10% separating polyacrylamide gel. Using a 1 mL micropipette, a freshly prepared solution was poured into the glass plate chamber in such a way as to avoid creating bubbles. The separating gel's edge was located less than 5 centimeters from the glass plate. For 45-60 minutes, it wanted to sit undisturbed so that the polymer could solidify.

The upper gel buffer and the other ingredients (Table 4.2.1) were mixed in different proportions to make a 4 % stacking gel, and the solution was poured using a 1 mL micropipette to prevent bubble formation. For polymerization and the generation of well-loaded protein samples, a comb was carefully positioned in the cleft of two glass plates and let to settle for thirty to forty-five minutes. After polymerizing the polyacrylamide gel, the comb was taken out, and the gel with the casting glass plate was set onto a vertical electrophoresis tank. Protein samples were treated in 2X sample loading buffer and heated at 95°C for 5 minutes to denature them before being loaded onto the gel. Each heat-treated denatured protein sample was mixed with 2 µL of a 0.1% bromophenol blue solution and loaded into the well. 5 µL of protein marker (Benchmark, Invitrogen, or Pre-stained protein Marker, NEB, England) was loaded and run alongside the samples to find out the desired protein size.

SDS running buffer was used for electrophoresis and was carried out at 100 volts until the tracking dye reached the bottom line of the gel. Once the electrophoresis was finished, the glass plates were submerged in distilled water to remove the plate and separate the gel.

**Table 4.2.1:** Polyacrylamide gel composition

| Ingredients | Separating gel 10% | Stacking gel 4% |
|---|---|---|
| dH20 | 1.9 mL | 1.5 mL |
| Lower or separating gel buffer | 1.3 mL | - |
| Upper or stacking gel buffer | - | 0.625 mL |
| 305polyacrylamide | 1.7 mL | 0.325 mL |
| 10%APS | 50 µL | 12.5 µL |
| TEMED | 2.0 µL | 2.0 µL |

#### 4.2.2.9 Staining and destaining of the PAGE gel

The glass plates were put in distilled water, and the gel was taken off the glass plates with the plate separator. After that, the gel was stained for 30 minutes in a rotary shaker containing 0.2% Coomassie blue staining solution (2 g of Coomassie Brilliant Blue G-250, 450 mL of methanol, 450 mL of $dH_2O$, and 100 mL of glacial acetic acid). The gel was then transferred to a container containing a destaining solution (50 mL of methanol, 70 mL of glacial acetic acid, and 880 mL of $dH_2O$) and thoroughly washed several times to remove the staining dye. After being washed 4-5 times, the gel was kept in $dH_2O$ overnight in a rotary shaker to reveal the clear protein bands.

#### 4.2.2.10 Protein band observation and molecular weight determination from SDS-PAGE

Protein bands of the acrylamide gel were observed at 302 nm UV illumination by a gel documentation system (Gel Doc, Bio-Science Corporation), and snaps were taken by the attached camera of the gel documentation system (Alpha Imager Mini, USA, Bio-Science Corporation). The molecular weight (kDa) of each protein was determined by calculating the relative migration distance of the desired protein band with the marker protein ladder. The migration distance of each of the protein bands was calculated using ImageJ (ij150-win-jre6), a Java-based image processing open-source program developed at the National Institutes of Health and the Laboratory for Optical and Computational Instrumentation.

#### 4.2.2.11 Anticancer activity test on HeLa cancer cell line

Human cervical cancer cell line (HeLa) was kept in DMEM (Dulbecco's Modified Eagles' Medium) with 1% penicillin-streptomycin (1:1), 0.2% gentamycin, and 10% fetal bovine serum (FBS). HeLa cells (2×104/100 µL) were seeded separately onto a 96-well plate and incubated at 37 ºC with 5% $CO_2$. Cytotoxicity was tested for proteinase K digested cry proteins to determine their anticancer activity. After 24 hours of cell line culture, 25 µL of filter-sterilized proteinase K digested parasporal protein from each Bt strain was added into each well. In such a manner, 25 µL of filter-sterilized proteinase K digested parasporal protein from each Bt strain was added into each well. Anticancer activity was examined after 24-48 hours of incubation of the proteolytically digested parasporal protein at 37 ºC under an inverted light microscope (Optika Italy, XDS-2 ERGO) at 40X magnification. Duplicate wells were used for each sample, and the mortality of HeLa cells was determined by qualitative assessment in comparison with HeLa cells without any solvent.

#### 4.2.2.12 Parasporal inclusions observation under phase-contrast microscope

Nonhemolytic Bt strains from a pure culture of LB agar medium were inoculated in 10 mL T3 broth and incubated at 30 °C in an orbital shaking incubator (Excella E25, New Brunswick Scientific) for approximately 7 days to allow sporulation. Then, 10 µl of sporulated cultures from each strain were smeared on the glass slides. After covering the glass slide with a coverslip each sample was observed under a phase-contrast microscope (Primo Star, ZEISS) at 100x (object) magnification. Strains with juxtaposed glowing spores and dark crystal protein were considered as Bt according to the method described by Bernhard *et al.,* (1997).

## 4.3 Results

### 4.3.1 Parasporal protein band observation from SDS-PAGE

The SDS PAGE was performed for all of the 98 nonhemolytic Bt strains to identify their parasporal protein band, and among them, only 7 strains MyIa2, FHSc4, KkSc1, RaSd1, MyIb1, MyIa1, and SgSp2 were found to produce one or more parasporal proteins having varied sizes (17 kDa to 78 kDa). During SDS-PAGE analysis, obtained protein concentrations of these 7 Bt strains are shown in Table 4.3.1, and their SDS-PAGE derived protein bands are shown in Figure 4.3.1. The obtained protein bands from different Bt strains are listed in a tabular form (Table 4.3.2).

**Table 4.3.1:** Absorbance at 595nm and concentrations of parasporal proteins.

| Bt strains | Absorbance at 595 nm | Concentration (mg/mL) |
|---|---|---|
| **MyIb1** | 0.315 | 0.561 |
| **RaSd1** | 0121 | 0.214 |
| **FHSc4** | 0.104 | 0.180 |
| **KkSc1** | 0.196 | 0.349 |
| **MyIa1** | 0.090 | 0.160 |
| **SgSp2** | 0.338 | 0.589 |
| **MyIa2** | 0.227 | 0.404 |

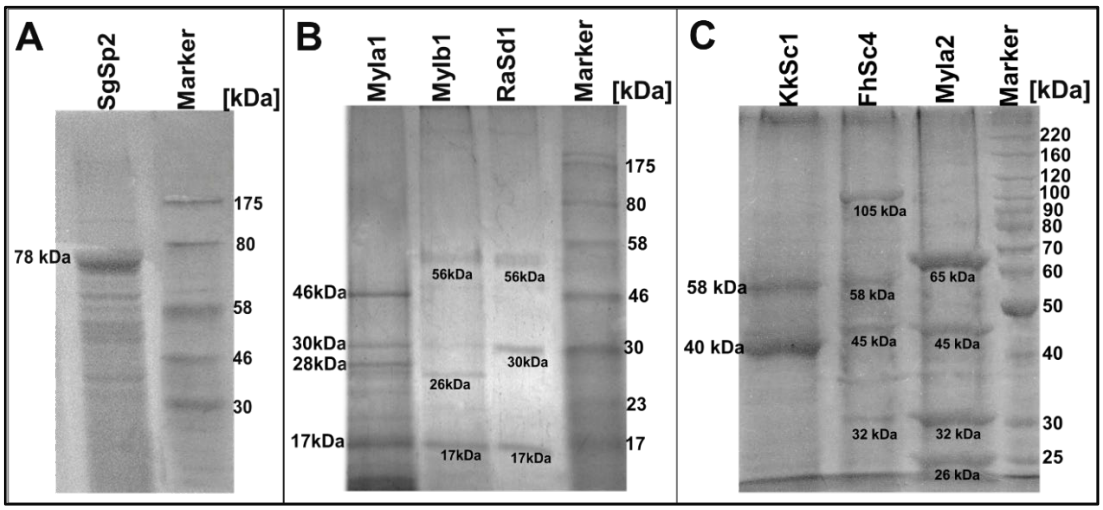



**Figure 4.3.1:** SDS-PAGE of solubilized parasporal inclusion protein of indigenous nonhemolytic *B. thuringiensis* strains (A-C).

**Table 4.3.2:** List of nonhemolytic Bt strains and their parasporal proteins.

| Name of *B. thuringiensis* strains | Parasporal protein band size (kDa) |
|---|---|
| **FHSc4** | 105, 58, 45, 32 |
| **KkSc1** | 58, 40 |
| **MyIa1** | 46, 30, 28, 17 |
| **MyIa2** | 65, 45, 32, 26 |
| **MyIb1** | 56, 28 |
| **RaSd1** | 56, 30 |
| **SgSp2** | 78 |

So far, researchers have identified six families of parasporin (PS1–PS6), which consist of 19 different parasporins. These parasporins are produced by at least 11 strains of *B. thuringiensis* (Bt). The majority of these discoveries have been made in three countries: Japan, Vietnam and Canada. PSs can be categorized into two groups based on their molecular mass. The first group includes PS1, PS3, and PS6, which have a higher molecular mass of around 80 kDa. These PSs are converted into active 60 kDa molecules through a proteolytic process. The second group consists of PS2, PS4, and PS5, which have a lower molecular mass ranging from 30 to 37 kDa. These PSs are derived from precursors and are transformed into 30 kDa molecules through proteolytic cleavage. In this study, some nonhemolytic Bt strains produced the prominent protein bands on the SDS gel closely matched with the known parasporins, PS1, PS2, PS4, and PS5. About 78 kDa parasporal protein was found from the SgSp2, which was close to the 81 kDa PS1. Besides, a 28 kDa protein band was found from MyIa1, MyIb1, and a 30 kDa band was found from MyIa1, RaSd1, which was close to the 30kDa of PS2Aa2. Besides, a 32 kDa protein band was found from FHSc4, MyIa2, which was close to the 30 kDa of PS4Aa1, and the 33 kDa of PS2Ab1, and PS5Aa1. These findings indicated the possible presence of PS1, PS2, PS4, and PS5.

Moreover, protein bands around 56 kDa to 58kDa were visible from RaSd1, MyIb1, FHSc4, and KkSc1 strains; and protein bands around 40 kDa to 46 kDa were visualized from MyIa1, MyIa2, FHSc4 and KkSc1 strains. These, along with a few other unpredicted-sized proteins, were under investigation to find novel parasporins.

### 4.3.2 Anticancer potential in HeLa cell line

Proteinase K digested solubilized parasporal protein of each of these 7 indigenous nonhemolytic *B. thuringiensis* strains were tested to detect their cytotoxicity to HeLa cell lines. Concentrations and doses of parasporal proteins used in the HeLa cell lines are shown in Table 4.3.3. Although a 78kDa parasporal protein that was nearly similar to the 81 kDa PS1Ac2 protein was obtained from SgSp2, no cytotoxic effect on the HeLa cell line was observed with the proteinase K treated protein from the said strain. Besides, among 7 Bt strains, proteinase K-treated solubilized parasporal proteins of 3 strains, MyIa2, FHSc4, and KkSc1 were found to be cytotoxic to the HeLa cell line (Figure 4.3.2).

**Table 4.3.3:** Concentrations and dose of parasporal proteins used in the HeLa cell line.

| Name of Bt strains | Protein Concentration (mg/mL) | Protein Concentration (µg/mL) | Protein added in each well (µg) | The final concentration of protein (µg/µL) |
|---|---|---|---|---|
| **MyIb1** | 0.561 | 561 | 14.025 | 0.2805 |
| **RaSd1** | 0.214 | 214 | 5.35 | 0.107 |
| **FHSc4** | 0.18 | 180 | 4.5 | 0.09 |
| **KkSc1** | 0.349 | 349 | 8.725 | 0.1745 |
| **MyIa1** | 0.16 | 160 | 4 | 0.08 |
| **SgSp2** | 0.589 | 589 | 14.725 | 0.2945 |
| **MyIa2** | 0.404 | 404 | 10.1 | 0.202 |

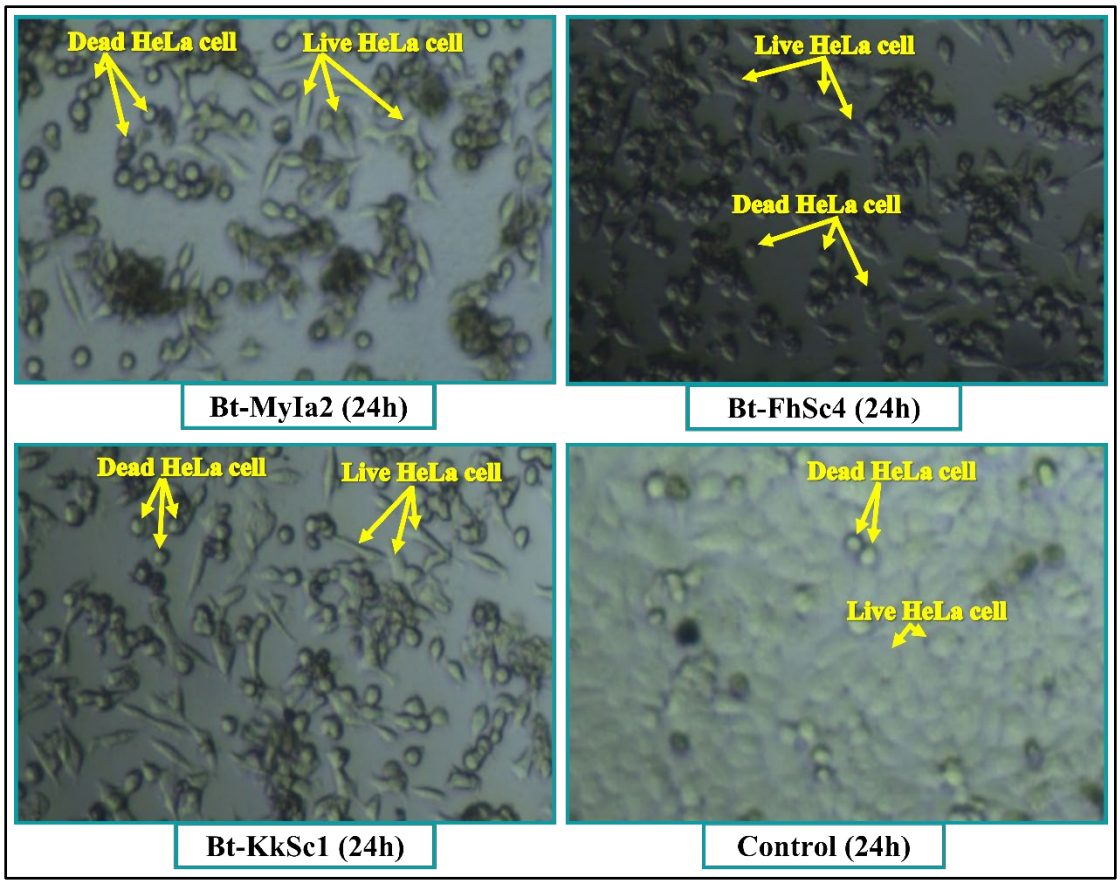



**Figure 4.3.2:** Proteinase K digested solubilized parasporal inclusion protein of 3 indigenous nonhemolytic Bt strains showed cytotoxicity against the HeLa cell line after 24 hours of incubation.

Unexpectedly, none of the putative parasporal proteins obtained from the 7 Bt strains produced any visible protein fragments on SDS-PAGE upon proteinase K treatment (Figure 4.3.3). This finding contradicts the existence of six parasporin families (PS1-PS6). This indicated that these parasporal proteins belonged to either a non-δ-endotoxin family, a non-Cry family, or an unidentified class of proteins. The parasporal protein fragments digested by proteinase K were likely too small to be measured using SDS-PAGE. As a result, the protein bands from the proteinase K digested parasporal protein fragments of the indigenous nonhemolytic Bt strains were not evident in the SDS gel. Although proteinase K digested solubilized parasporal protein of 3 Bt strains MyIa2, FHSc4, and KkSc1 showed cytotoxicity to the HeLa cell lines, they did not produce any visible protein bands in the SDS gel. Therefore, any further experiments on cytotoxicity were not performed using these parasporal proteins.

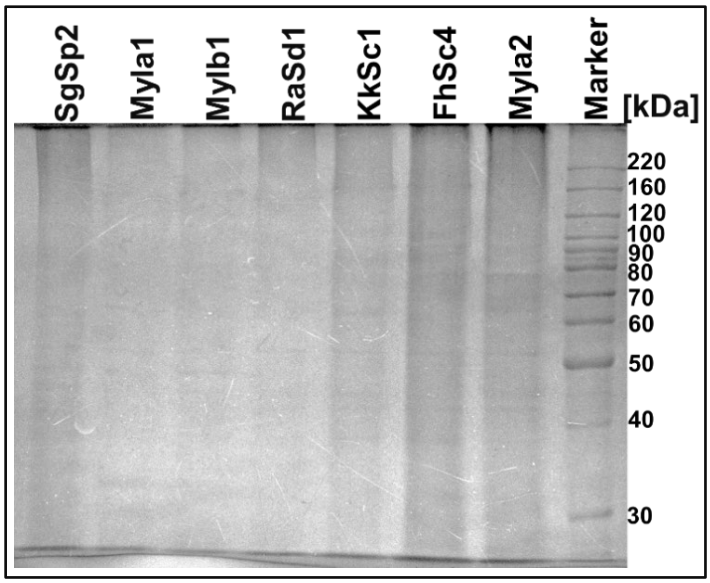



**Figure 4.3.3:** SDS PAGE of proteinase K treated solubilized parasporal inclusion protein of indigenous nonhemolytic *B. thuringiensis* strains.

### 4.3.3 Parasporal inclusion protein observation under phase-contrast microscopy

Among 98 indigenous nonhemolytic Bt strains, 10% n-hexane purified, HeLa-cell-killing parasporal inclusion bodies from the 3 indigenous nonhemolytic *B. thuringiensis* strains, MyIa2, FHSc4, KkSc1, along with the reference Bt strain HD-73 were observed under phase-contrast microscope. Purified parasporal inclusion bodies under phase-contrast microscope were revealed as round in KkSc1, cubic and bipyramidal shape in reference strain HD73, and indigenous MyIa2 and FHSc4 (Figure 4.3.4). Microscopic observation of parasporal inclusions supports the previously observed protein band from SDS-PAGE of indigenous Bt strains as parasporal protein.

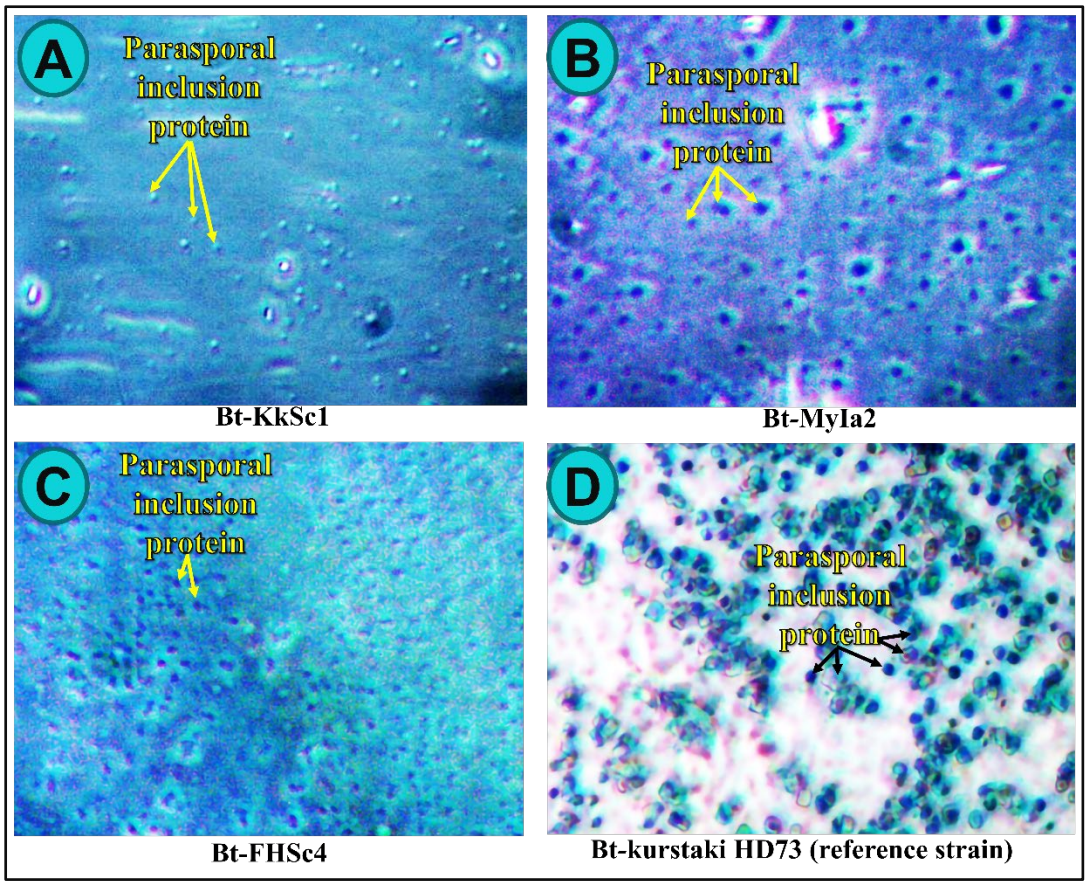



**Figure 4.3.4:** n-hexane (10%) purified parasporal inclusion protein under phase-contrast microscope at 100X. (A-C) indigenous Bt strains, (D) reference Bt strain HD-73.

The main goal of this research was to find the parasporin protein that can suppress cancer cells from native nonhemolytic *B. thuringiensis*. Through this study, we have identified parasporal proteins producing indigenous nonhemolytic *B. thuringiensis* strains. The presence and size of all these parasporal proteins were detected by SDS-PAGE, and parasporal inclusions were observed under a phase-contrast microscope. Although some of the proteins we identified have similar sizes to PS1, PS2, PS4, and PS5, none of these parasporal proteins can be called parasporins or any other cry proteins because they are completely digested by proteinase K. Not being one of the parasporins or cry proteins, the presence of cytocidal toxicity against the human cancer cell line of the proteins under study directs the research to investigate their unknown biological identity.

# CHAPTER 5:

## Characterization of parasporal anticancer proteins from indigenous hemolytic non-insecticidal *Bacillus thuringiensis* strains

# CHAPTER 5: Characterization of parasporal anticancer proteins from indigenous hemolytic non-insecticidal *Bacillus thuringiensis* strains

## 5.1 Introduction

The parasporal inclusion body is a crystalline protein complex found outside a spore in some bacteria. It was found from the literature that *Bacillus thuringiensis* (Travis & Glare, 2000), *B. sphaericus* (Davidson *et al.*, 1981; Payne *et al.*, 1984), *B. laterosporus* (Montaldi & Roth, 1990; Smirnova *et al.*, 1996) and *B. subtilis* (Rubikas *et al.*, 1987) produces parasporal crystals next to their endospores during spore formation. Depending on the amino acids in the protein, the parasporal crystals can be bipyramidal, cuboidal, rhomboid, spherical, or a combination of two or more different types of crystals. Like parasporal Cry, *B. thuringiensis* also produced another kind of parasporal inclusion protein, called surface layer or S-layer protein or simply SLP; mostly round or oval-shaped (Sun *et al.*, 2001). Like Cry and Cyt proteins, S-layer proteins of *Lysinibacillus sphaericus* have mosquitocidal activity (Thanabalu *et al.,* 1991, Allievi *et al.,* 2014). Besides, SLPs of *B. thuringiensis* had selective toxicity to coleopteran beetles (Peña *et al.*, 2006), cattle ticks (Lormendez *et al.,* 2019),

In addition to the insecticidal activity, some Cry proteins of *B. thuringiensis* have anti-protozoan (Kondo *et al.*, 2002), and some have anticancer activity (Kim *et al.*, 2000; Mizuki *et al.*, 2000; Nagamatsu *et al.*, 2010). The anticancer Cry proteins of *B. thuringiensis* are known as parasporins (PS), and most of them have potential cytocidal activity against various human carcinoma cells (Gonzalez *et al.*, 2011; Jung *et al.*, 2007; Lenina *et al.*, 2014; Mizuki *et al.*, 1999; Nagamatsu *et al.*, 2010; Ohba *et al.*, 2009; Poornima *et al.*, 2010; Uemori *et al.*, 2007; Wong., 2010; Yasutake *et al.*, 2006). The carcinoma cell-killing property of PS has raised new hope for cancer treatment globally. After the invention of parasporins, several studies have shown that these novel

parasporal proteins (parasporins) are nonhemolytic and non-insecticidal but have selective anticancer activity on different cancer cell lines (Ohba *et al.*, 2009). Besides the anticancer Cry protein, the parasporal SLP of *B. thuringiensis* also had human cancer cell-killing activity (Rubio *et al.,* 2017).

The proteinase K or trypsin-digested solubilized parasporal protein of *B. thuringiensis* exhibits a specific cytotoxic effect on cancer cells (Mizuki *et al.,* 2000; Brasseur *et al.,* 2015), which could be significant in developing effective protein-based anticancer treatments. Therefore, a promising opportunity arises to develop protein-based therapies for cancer treatment by utilizing the anticancer properties of parasporal inclusion proteins derived from Bt. Reports are also found on anticancer Cry proteins PS2, PS3, and PS4 isolated from hemolytic *B. thuringiensis* strains (Ichikawa *et al.*, 2008). This study was thus conducted to discover the novel parasporal anticancer proteins from previously identified twenty-three indigenous non-insecticidal hemolytic *B. thuringiensis* strains. Each of twenty-three *B. thuringiensis* was sporulated entirely and then the competence of these strains to synthesize parasporal proteins were assessed using SDS-PAGE. Five indigenous Bt strains were identified to produce one or more parasporal proteins. The cytotoxic effects of the solubilized crude, proteinase K, and trypsin-digested parasporal proteins were investigated on the HeLa and Vero cell lines. Their hemolytic as well as insecticidal activity was also observed. Additionally, morphological characteristics of parasporal inclusions were examined to enhance comprehension of their inherent properties.

## 5.2 Materials and Methods

### 5.2.1 Materials

#### 5.2.1.1 Chemicals

Acrylamide ($C_3H_5NO$), agar ($C_{14}H_{24}O_9$), ammonium persulfate (APS) {$(NH_4)_2S_2O_8$}, n,n′-methylenebisacrylamide $CH_2[NHC(O)CH=CH_2]_2$, bromophenol blue ($C_{19}H_{10}Br_4O_5S$), coomassie blue G-250 ($C_{47}H_{48}N_3NaO_7S_2$), $dH_2O$, DTT (dithiothreitol) ($C_4H_{10}O_2S_2$), glacial acetic acid ($CH_3COOH$), glycine ($C_2H_5NO_2$), phenylmethylsulfonyl fluoride (PMSF) ($C_7H_7FO_2S$), $Na_2EDTA$ ($C_{10}H_{14}N_2O_8.2Na.2H_2O$), ethanol ($CH_3CH_2OH$), n-hexane ($C_6H_{14}$), glycerol, HCl, methanol ($CH_3OH$), $MnCl_2$, KCl, NaCl, $K_2HPO_4$, $Na_2HPO4$, $KH_2PO4$, $Na_2CO_3$, phosphoric acid ($H_3PO_4$), SDS {$CH_3(CH_2)_{11}OSO_3Na$}, tris-base ($C_4H_{11}NO_3$). The chemicals used in this study were of molecular grade, and their lists and company names are mentioned in **Appendix III**.

#### 5.2.1.2 Equipment

Autoclave machine, biosafety cabinet, centrifuge machine, electronic balance, microbiological incubator, magnetic stirrer, micropipettes, orbital shaker incubator, pH meter, phase-contrast microscope, SDS gel electrophoresis power supply, -20 $^oC$ fridge, 4 $^oC$ refrigerator, sonication machine, spectrophotometer, thermo stated shaking water bath, vortex mixture, etc., apparatuses used in this study are mentioned in the respective methods section, and their company and models are mentioned in **Appendix IV**

#### 5.2.1.3 Media

Luria Bertani (LB) agar and broth, and T3 broth, were used in this study, and their compositions are mentioned in **Appendix I**.

#### 5.2.1.4 Reagents and Solutions

Bacto-tryptone, bacto-tryptose, bromophenol blue (1%), yeast extract, proteinase K, phosphate-buffer, phosphate-buffered saline (PBS), normal saline, Tris-HCl, EDTA, parasporal protein solubilizing buffer (pH 11.0), Bradford reagent, proteinase K dissolving solution, 0.1M PMSF, 30% acrylamide-bisacrylamide, 10% APS, SDS upper gel buffer (pH 6.8), SDS lower gel buffer (pH 8.8), SDS electrophoresis buffer, SDS staining and destaining solution, protein marker, protein sample preparation buffer, etc.,

reagents and solutions used in these experimental procedures were mentioned in the respective methods section. Their composition and preparation methods are mentioned in **Appendix II.**

#### 5.2.1.5 Bacterial strains

A total of 23 β-hemolytic non-pesticidal *B. thuringiensis* strains **(25L, 34L, 45L, 19S, 28S, 55S, 57S, BD59S, 1i, Dse1, DSN4, DSN7, Dsf7, Dsh4, JeSa1, JSa3, Leaf-28, Leaf-31, Leaf-58, SSa3, Soil-46, Soil-57, and Soil-108)** were studied.

### 5.2.2 Methods

#### 5.2.2.1 Identification of Bt strains by the chromogenic differential method

Initially, all of these indigenous non-insecticidal 23 Bt strains were screened by a rapid and easy chromogenic *Bacillus* agar identification and isolation technique. This agar medium contains peptic digests of animal tissues, meat extracts, mannitol, and phenol red. *B. cereus* has the enzyme beta-glucosidase, which cleaves the chromogenic mixture in the medium, resulting in blue colonies. Since *B. cereus* and *B. thuringiensis* are biochemically equivalent, both species will develop as blue-green colonies on this medium; however, while *B. cereus* displays flat colonies with identifiable blue cores, and *B. thuringiensis* displays irregular borders.

#### 5.2.2.2 Maintenance of Culture

Strains were maintained in LB agar by subculturing and stored at 4 ºC for further work.

#### 5.2.2.3 Parasporal crystal protein preparation

As mentioned in the previous chapter in section 4.2.2.1

#### 5.2.2.4 Spore crystal separation using a simple sonication method

In this method, indigenous pure Bt culture was added to 50 mL of T3 broth medium and incubated at 30 ºC for 7 days at 120 rpm on an orbital shaker. Then collect all the spore crystals from the culture in a 50 mL Falcon tube by two to three times centrifugation at 3300 × g for 10 minutes (TOMY MX-305 high speed refrigerated micro centrifuge, Japan). The supernatant was discarded, and the pellet was washed

with ice-cold $dH_2O$. This step was repeated three times. The pellet was resuspended in 15 mL $dH_2O$ and sonicated at 100 W for 30 minutes with a 10-second pulse (Omni-Ruptor 4000 Ultrasonic Homogenizer, OMNI International, USA). Then the pellet was centrifuged at 5800 × g for 5 minutes, collected, and washed twice with cold water. As cry proteins are light-sensitive (Pozsgay *et al.*, 1987), all of the Falcon tubes were wrapped with aluminium foil. A detailed workflow of spore crystal separation using a simple sonication method is shown in Figure 5.2.1.

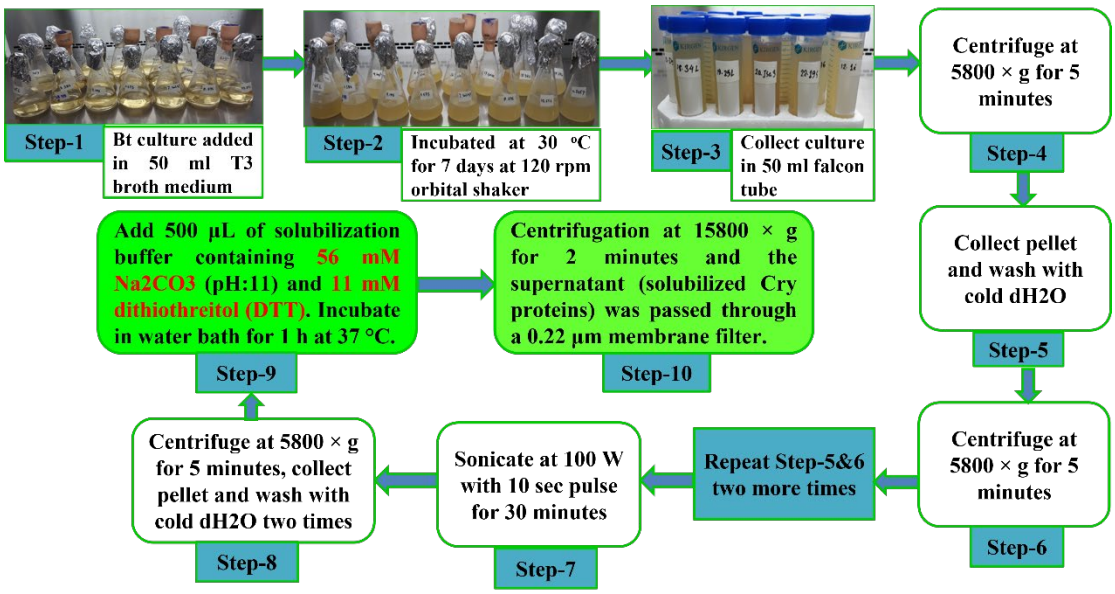



**Figure 5.2.1:** Workflow of spore crystal separation using simple sonication method

#### 5.2.2.5 Alkali solubilization of parasporal Cry proteins

As mentioned in the previous chapter in section 4.2.2.4

#### 5.2.2.6 Determination of cry protein concentration by the Bradford Assay

As mentioned in the previous chapter in section 4.2.2.5

#### 5.2.2.7 Proteolytic digestion, activation, and sterilization of parasporal protein proteins

Each of the solubilized cry protein samples was digested separately with two different proteolytic enzymes, proteinase K (final concentration of 185 µg/mL) and trypsin (final concentration of 300 µg/mL) by incubating for about 1 hour at 37 °C in a water bath (equipment name). After that phenylmethylsulfonyl fluoride (PMSF) was added (final concentration: 1 mM) to each cry protein sample to stop the proteolytic processing.

Then each sample was sterilized with a 0.22 µm membrane filter. All of the microcentrifuge tubes containing parasporal proteins were wrapped with aluminium foil to protect them from light.

#### 5.2.2.8 SDS-PAGE analysis

As mentioned in the previous chapter in section 4.2.2.7

##### 5.2.2.8.1 Preparation of SDS polyacrylamide gel

As mentioned in the previous chapter in section 4.2.2.8

##### 5.2.2.8.2 Staining and destaining of the PAGE gel

As mentioned in the previous chapter in section 4.2.2.9

#### 5.2.2.9 Protein band observation and molecular weight determination from SDS-PAGE

As mentioned in the previous chapter in section 4.2.2.10

#### 5.2.2.10 Cytotoxic effects of the solubilized parasporal protein on Artemia

To do this experiment, a small number of cysts were poured into 1 L of artificial seawater (15.15 g NaCl, 2.98 g $MgCl_2 \cdot 6H_2O$, 2.86 g $MgSO_4 \cdot 7H_2O$, 0.65 g $CaCl_2$, 0.16 g $NaHCO_3$, 0.414 g KCl) in a conical flask and two bright LED lamp was lighted and placed in front of it and kept light until they hatched. After 48 hours, newly hatched brine shrimp nauplii ($n$ = 10/sample/concentration) were transferred to 5 mL sea water containing test-tube. The effects of Cry proteins at different concentrations ranging from 68 to 269 µg/mL for the survival of *Artemia* nauplii were recorded. Negative controls were prepared using the seawater containing BSA at 1000 µg/mL. After 24 h, the dead nauplii were counted from each concentration of the cry protein sample and plotted on a graph to correlate mortality and sample concentration. Finally, probit analysis was performed to determine the median lethal concentration (LC50) at which 50% of the study population was killed.

#### 5.2.2.11 Hemolysis assays

Sheep erythrocytes were separated from whole blood by adding sterile glass beads, and after continuous shaking, red blood cells were separated. Red blood cells containing blood serum were added to the blood agar, and the blood agar plate was prepared and checked for any contamination at 37 °C for 18 h. Subsequently, the Cry protein samples were incubated at 37 °C in each blood agar plate for 18 h and observed.

#### 5.2.2.12 Bioassay

##### 5.2.2.12.1 Insect larvae rearing

The larvae of *Bactrocera cucurbitae* were fed a semi-liquid artificial diet that was manufactured in-house. The adults were housed in a netted stainless-steel cage (100 cm × 100 cm × 70 cm). The insects were usually supplied with a laboratory diet (yeast extract, casein, and sugar, 1: 1: 2) and water-soaked cotton. About 200–250 mature fruit flies were kept in a stock cage. The rearing room was kept at a constant temperature of 28 °C and relative humidity (RH) of 70–80%. The matured flies in the cage were given a piece of the sweet gourd for oviposition to collect a large number of eggs. After 2 hours, the slice of sweet gourd was taken from the adult cage and put in a plastic dish with sawdust for additional larval growth.

##### 5.2.2.12.2 Insecticidal activity test

The toxicity of the parasporal inclusion proteins of Bt strains was analyzed *in-vivo* by bioassay against the 3rd instar *B. cucurbitae* larvae. Both trypsin-treated and untreated Cry proteins were prepared and mixed with 10 gm of boiled and smashed sweet gourd paste, on which 20 larvae were kept and fed at 25-28 °C and 70–80% RH, with a photoperiod of 16:8 (L:D). The mortality was then scored for the cry proteins along with a parallel control prepared with sterile distilled water. The concentration of cryoglobulins in the suspension was determined by the Bradford method. In every instance, bioassays were carried out in triplicate.

#### 5.2.2.13 Cytotoxicity test on HeLa cell line

As mentioned in the previous chapter in section 4.2.2.11

#### 5.2.2.14 Cytotoxicity test on Vero cell line

Monkey kidney epithelial cell line (Vero) was grown in DMEM (Dulbecco's Modified Eagles' Medium) with 1% penicillin-streptomycin (1:1), 0.2% gentamycin, and 10% fetal bovine serum (FBS). Vero cells (2×104/100 µL) were seeded separately onto a 96-well plate and incubated at 37 ºC with 5% $CO_2$. Cytotoxicity was tested separately for both trypsin and proteinase K-digested cry proteins to determine whether trypsin-activated or proteinase K-activated cry proteins had shown cytotoxicity. After 24 hours of cell line culture, 25 µL of filter-sterilized trypsin digested cry protein from each Bt strain was added into each well. In such a manner, 25 µL of filter-sterilized proteinase K digested cry protein from each Bt strain was added into each well. Cytotoxicity was examined after 24-48 hours of incubation of the proteolytically digested cry at 37 ºC under an inverted light microscope (Optika Italy, XDS-2 ERGO) at 40X magnification. Duplicate wells were used for each sample, and the viability of Vero cells was determined by qualitative assessment in comparison with Vero cells without any solvent.

#### 5.2.2.15 Morphology analysis of parasporal crystals

Strains were characterized based on the shape and size of parasporal crystals by light-phase microscopy as well as scanning electron microscopy.

##### 5.2.2.15.1 Parasporal inclusion protein observation under the light phase microscopy

To observe the parasporal crystal protein of indigenous *B. thuringiensis* strains, a previously described method (Rampersad *et al.,* 2002) was applied with a slight modification. A brief narration of the method was, that 10 µL of sonicated spore cry mixture was mixed with 10 µL of PBS on a clean glass slide and smeared. Then air-dry the sample and heat-fix it by running the opposite side of the slide above a mild stage of fire flame. Then, the slide was merged into 0.2% Coomassie stain (1 gm of Coomassie brilliant blue G-250, 225 mL of ethanol, 225 mL of $dH_2O$, and 50 mL of glacial acetic acid) within a Coplin staining jar for 3 minutes. Then the slide was washed with $dH_2O$ for 1 minute. The slide was then merged into the destaining solution for 3 minutes. Then, the slide was air-dried and observed under the light phase of a phase-contrast microscope at 100X (object) magnification (Primo Star, ZEISS).

#### 5.2.2.15.2 Parasporal inclusion protein observation through scanning electron microscope (SEM)

Sonicated and $dH_2O$-washed spore crystal mixture (mentioned in 5.2.2.4) of indigenous *B. thuringiensis* strains of interest were washed three times with ice-cold 95% ethanol and dried by air in a microcentrifuge tube. Then, 500 µL of 95% ethanol was added to the air-dried spore crystal mixture and mixed well by pipetting. It was then stored at 4 ºC in a refrigerator. The spore crystal mixture was subsequently mounted on aluminium and fastened. After that, an auto-fine coater was used to coat the samples (JEOL JFC-1600). The SEM observation was performed following the manufacturer's instructions using a 15 kV analytical scanning electron microscope (JEOL JSM-6490LA analytic SEM, Japan).

### 5.2.2.16 Biochemical characterization of parasporal protein synthesizing Bt strains

Biochemical characterization of parasporal anticancer protein synthesizing indigenous Bt strains was performed and recorded.

#### 5.2.2.16.1 Colony characteristics

Each strain was analyzed for its morphological and cultural characteristics. All strains were inoculated on LB agar plates and incubated overnight at 30 °C. Colony characteristics were observed and listed.

#### 5.2.2.16.2 Gram staining

For Gram staining, the cells were first fixed with heat to the glass slide and then stained with a basic dye, crystal violet. After which, the slides are mordanted with an I2-KI combination to set the stain, temporarily destained with 95% alcohol, and then counterstained with a lighter dye of a different hue (safranin). The slide was then observed under a bright-field microscope at 100X magnification.

#### 5.2.2.16.3 Starch hydrolysis test

Starch-nutrient agar plates were used to grow the strains at 30 ºC for 18 hours. After 18 hours at 30 ºC, the plates were flooded with 5-10 mL of Gram's iodine solution as

instructed in Bergey's Manual of Determinative Bacteriology. The starch in the medium is transformed into a blue-black hue when an iodine reagent reacts with it. Starch hydrolyzing bacteria, like *B. thuringiensis*, were inferred from isolations that exhibited a distinct zone surrounding their colonies.

#### 5.2.2.16.4 Voges-Proskauer (VP) test

When bacteria undergo sugar fermentation, they generate 2, 3-butanediol as a predominant final product, which then builds up in the surrounding medium. A volume of 3 mL of fresh culture was inoculated in a medium containing glucose, phosphate, and peptone. Subsequently, 0.2 mL of a 40% potassium hydroxide (KOH) solution and 0.6 mL of a 5% alpha-naphthol solution in pure ethanol were introduced. An intense shaking for 10-15 minutes resulting in a cherry-red color indicates a positive result, whereas a yellow-brown shade indicates a negative one.

#### 5.2.2.16.5 Indole production test

The Indole test was conducted by inoculating peptone water and incubating it at a temperature of 30 °C for 18-24 hours. Following incubation, 3–4 drops of Kovac's reagent were introduced to detect the formation of a crimson circular region at the surface of peptone water within a minute. The presence of a rich red hue signified the synthesis of indole from tryptophan.

#### 5.2.2.16.6 Methyl Red test

The methyl-red test was used to ascertain the capacity of microorganisms to oxidize glucose, resulting in the generation and stability of elevated levels of acidic end products. Next, a 2 mL aliquot of a recently cultivated bacterial culture in glucose phosphate medium, incubated for 18-24 hours at 30 °C, was transferred into a test tube. The addition of five drops of methyl red reagent was followed by quick mixing. A positive test will cause the methyl red indicator to become red at a pH level of 4.

#### 5.2.2.16.7 NaCl (7%) growth test

Strains were grown on a 7% NaCl solution at 30 °C overnight. The growth of the solution indicates a positive result.

#### 5.2.2.16.8 Gelatinase test

A gelatin hydrolysis test measures an organism's ability to produce gelatinase, a proteolytic enzyme that dissolves gelatin. The hydrolysis of gelatin indicates the presence of gelatinases. Bt strains were grown overnight on media containing gelatin and peptone at 30 °C. Since gelatin usually becomes liquid at temperatures of 28 ºC or higher, the tubes are placed in an ice bath or refrigerated at 4 ºC to ensure the result of gelatinase activity. The degree to which the medium melts is indicative of a positive result.

#### 5.2.2.16.9 Preservation of parasporal protein synthesizing Bt strains by glycerol stock

Bt strains that produced parasporal inclusions were streaked on LB agar plates and grown overnight at 30 ºC. Then, a single colony from each strain was grown overnight on LB broth at 37 °C at 160 rpm. Then 0.85 mL of culture was transferred to a sterile microcentrifuge tube containing 0.15 mL of sterile glycerol. The content of the centrifuge tube was mixed carefully by pipetting. A total of 10 sets of stocks were prepared for each of the strains and kept primarily at -20 ºC, and finally transferred to a -80 ºC freezer for long-term preservation of the strains.

## 5.3 Results

### 5.3.1 Identification of Bt strains by the chromogenic differential method

Among twenty-three (23) β-hemolytic but non-insecticidal *B. thuringiensis* strains **(25L, 34L, 45L, 19s, 28S, 55s, 57s, BD59S, 1i, Dse1, DSN4, DSN7, Dsf7, Dsh4, JeSa1, JSa3, Leaf-28, leaf-31, Leaf-58, SSa3, Soil-46, soil-57, and soil-108)** and three (3) reference *B. thuringiensis* strains (kurstaki HD-73, sotto, and japonensis), each were found to produce deep blue colonies in the chromogenic *Bacillus* agar plate (Figure 5.3.1) indicating their identity as *B. thuringiensis*. The centre of each colony was found as raised and margins were found as irregular as expected.

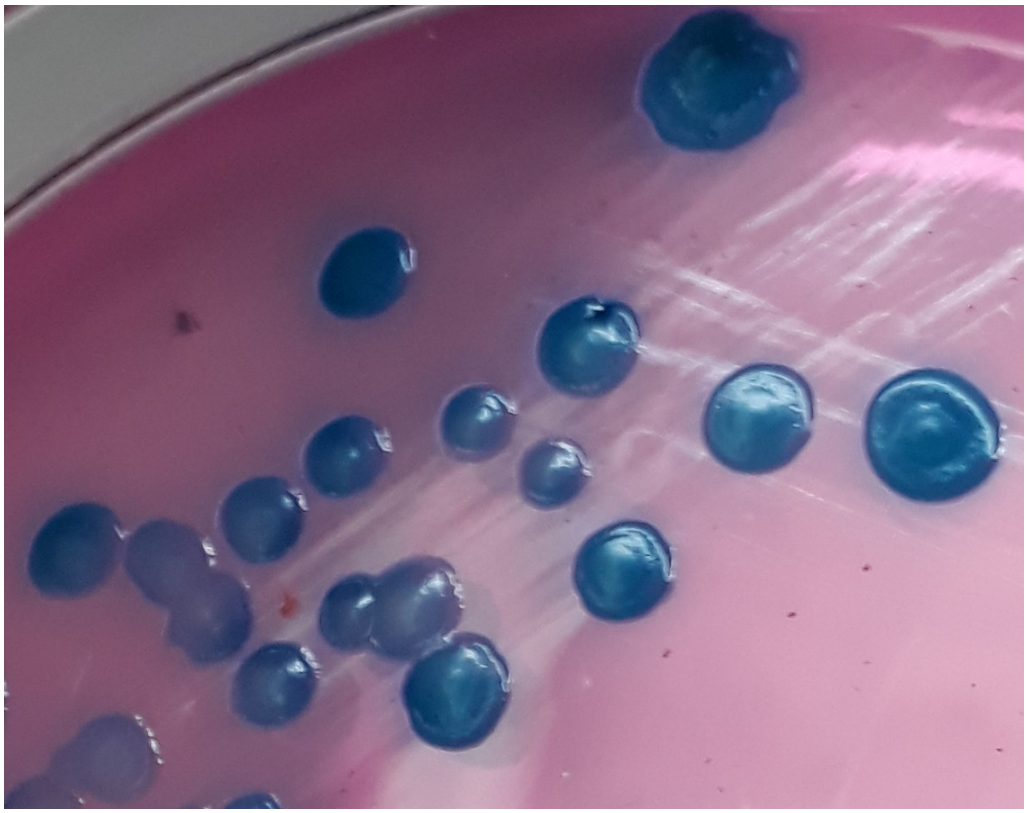


**Figure 5.3.1:** Close view of *B. thuringiensis* BD59S colonies on chromogenic agar plate

### 5.3.2 Parasporal protein band observation from SDS-PAGE

Twenty-three indigenous Bt strains (25L, 34L, 45L, 19s, 28S, 55s, 57s, BD59S, 1i, Dse1, DSN4, DSN7, Dsf7, Dsh4, JeSa1, JSa3, Leaf-28, leaf-31, Leaf-58, SSa3, Soil-46, soil-57, and soil-108) were screened for detection of their parasporal inclusion protein. Among the 23, only 5 strains, BD59S, Soil-46, 28S, 45L, and Dsh4 produced parasporal inclusion proteins as revealed on the SDS-PAGE (Figure 5.3.2), and 18 Bt

strains were found as acrystalliferous. The protein concentrations used in the SDS-PAGE were measured by the Bradford method (Table 5.3.1). Among the 5, Soil-46, and 28S were found to produce 4 parasporal proteins of identical molecular mass (~101 kDa, ~86 kDa, ~76 kDa, and ~28 kDa) (Figure 5.3.2. A). Besides, BD59S was also found to produce 3 different parasporal proteins of identical molecular mass (~101 kDa, ~86 kDa, and ~76 kDa) as isolated from Soil-46, and 28S (Figure 5.3.2. A). A single and identical ~103 kDa protein band in the SDS-PAGE (Figure 5.3.2. B) was identified from Dsh4 and 45L.

**Table 5.3.1:** Protein concentrations used in the SDS PAGE of Figure 5.3.2

| Bt strains | Protein concentration (mg/mL) | Protein band (kDa) |
|---|---|---|
| **45L** | 0.642 | 103 |
| **Dsh4** | 0.832 | 103 |
| **BD59S** | 1.016 | 101, 86, 76 |
| **28S** | 1.143 | 101, 86, 76 |
| **Soil-46** | 1.146 | 101, 86, 76 |

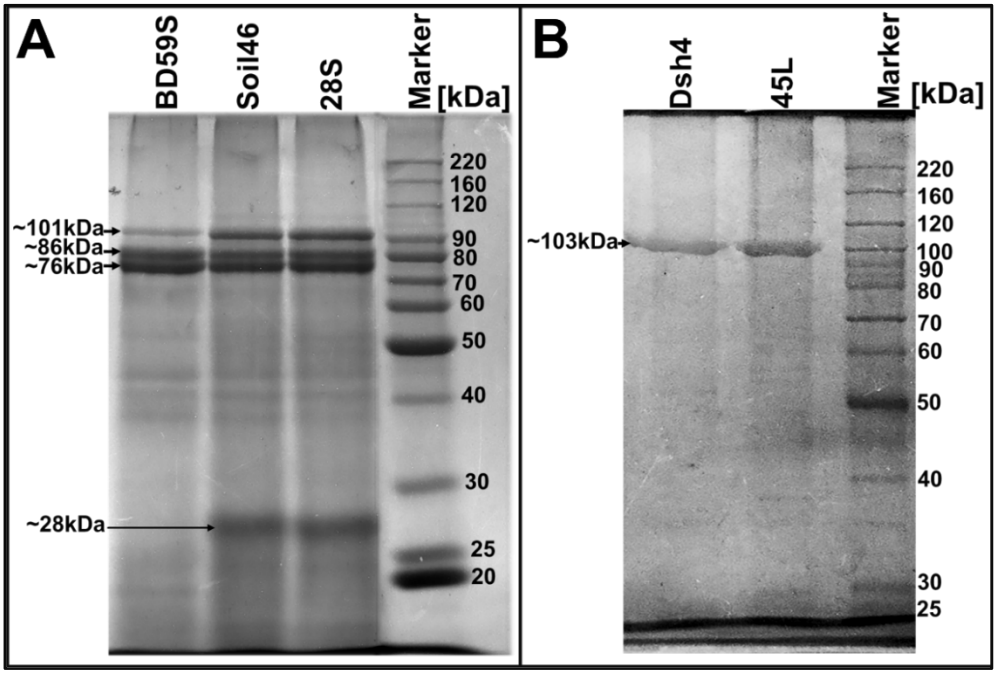



**Figure 5.3.2:** SDS PAGE of solubilized crude parasporal inclusion protein of indigenous *B. thuringiensis* strains. (A) Bt strains BD59S, 28S, and Soil-46, each strain was found to produce 3 different sizes ~101 kDa, ~86 kDa, and a ~76 kDa parasporal protein band and a 28 kDa band was observed to be produced from two strains Soil-46 and 28S; (B) Bt strains Dsh4 and 45L produced a single ~103 kDa protein band.

### 5.3.3 Trypsin digested protein band observation from SDS-PAGE

Solubilized parasporal proteins from each of five *B. thuringiensis* strains were digested with trypsin at 37 ºC for 1 hour. As a result, 101 kDa, 86 kDa and 76 kDa proteins of Bt strain 28S and Soil-46 produced five identical digested fragments, approximately 64, 51, 45, 41 and 32 kDa (Figure 5.3.3). Whereas, BD59S produced three digested fragments, approximately 64, 51 and 41 kDa. On the other hand, 103 kDa protein of Dsh4 and 45L, fragmented to a 45 kDa single band in the SDS-PAGE (Figure 5.3.3). Here, it is worth considering that, the 45 kDa fragment was absent only in BD59S while all four strains namely Dsh4, 45L, 28S and Soil-46 produced the similar sized upon digestion with trypsin.

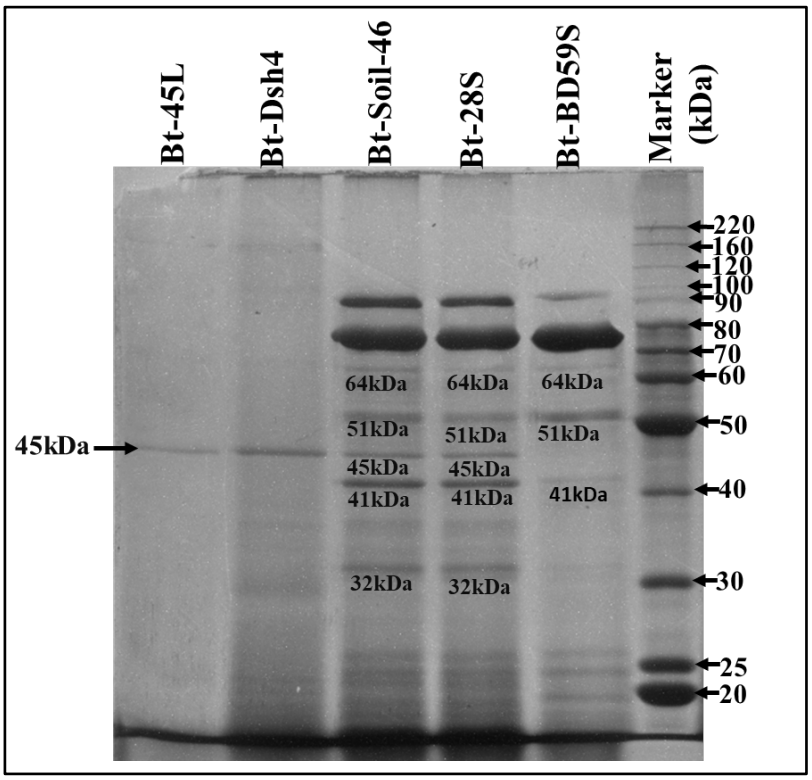



**Figure 5.3.3:** SDS PAGE of trypsin digested solubilized parasporal protein of BD59S, 28S, Soil-46, Dsh4, and 45L.

In this experiment, our obtained trypsin digested protein fragments did not match with any one of trypsin digested fragments of parasporin. Trypsin-digested PS1 is known to produce 55 to 56 kDa protein fragments, while trypsin-digested PS6 is known to produce a 59 kDa protein fragment.

### 5.3.4 Proteinase K digested protein band observation from SDS-PAGE

Each of the solubilized parasporal proteins from 5 indigenous *B. thuringiensis* strains were individually incubated with proteinase K in microcentrifuge tubes at 37 ºC for 1 hour. Then proteinase K digested parasporal proteins were observed in the SDS-PAGE. As a result of digestion, approximately 46 kDa digested single protein fragment was found in the SDS-PAGE from each of 5 strains (Figure 5.3.4).

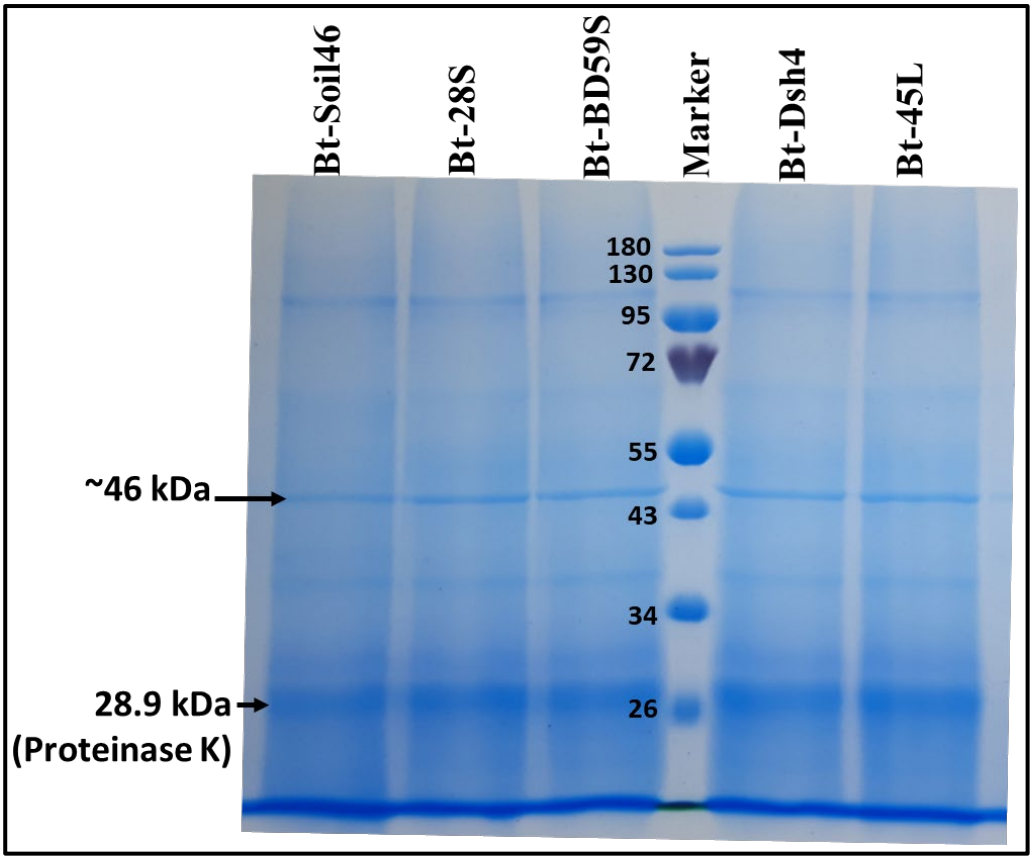



**Figure 5.3.4:** SDS PAGE of proteinase K treated BD59S, 28S, Soil-46, Dsh4, and 45L produced about 46 kDa protein band after 1 hour of incubation at 37ºC

### 5.3.5 Cytotoxic effects of the solubilized cry protein on Artemia

The solubilized and untreated crude parasporal inclusion proteins from five native *B. thuringiensis* strains (BD59S, Soil-46, 28S, 45L, and Dsh4) were analyzed to assess their *in vivo* cytotoxic impact (Table 5.3.2). The *in vivo* cytotoxicity test revealed that the brine shrimp nauplii (*Artemia salina*) were affected by the cytotoxicity, with an LC50 value of 114.192 µg/mL (Figure 5.3.5).

**Table 5.3.2:** Protein concentration and dose used in the *in vivo* cytotoxic experiment

| Name of Bt strains | Protein concentration added (mg/mL) | Final protein concentration (µg/mL) | LogC | No. of Artemia taken | No. of dead Artemia (12 h) | No. of Live Artemia (12 h) | Mortality (%) | Probit | LC50 µg/mL |
|---|---|---|---|---|---|---|---|---|---|
| **45L** | 0.208 | 68.64 | 1.837 | 20 | 6 | 14 | 30 | 4.48 | |
| **Dsh4** | 0.397 | 131.01 | 2.117 | 20 | 8 | 12 | 40 | 4.75 | |
| **28S** | 0.512 | 168.96 | 2.228 | 20 | 8 | 12 | 40 | 4.75 | **114.192** |
| **Soil-46** | 0.816 | 243.54 | 2.387 | 20 | 20 | 0 | 100 | 7.37 | |
| **BD59S** | 0.738 | 269.28 | 2.430 | 20 | 20 | 0 | 100 | 7.37 | |

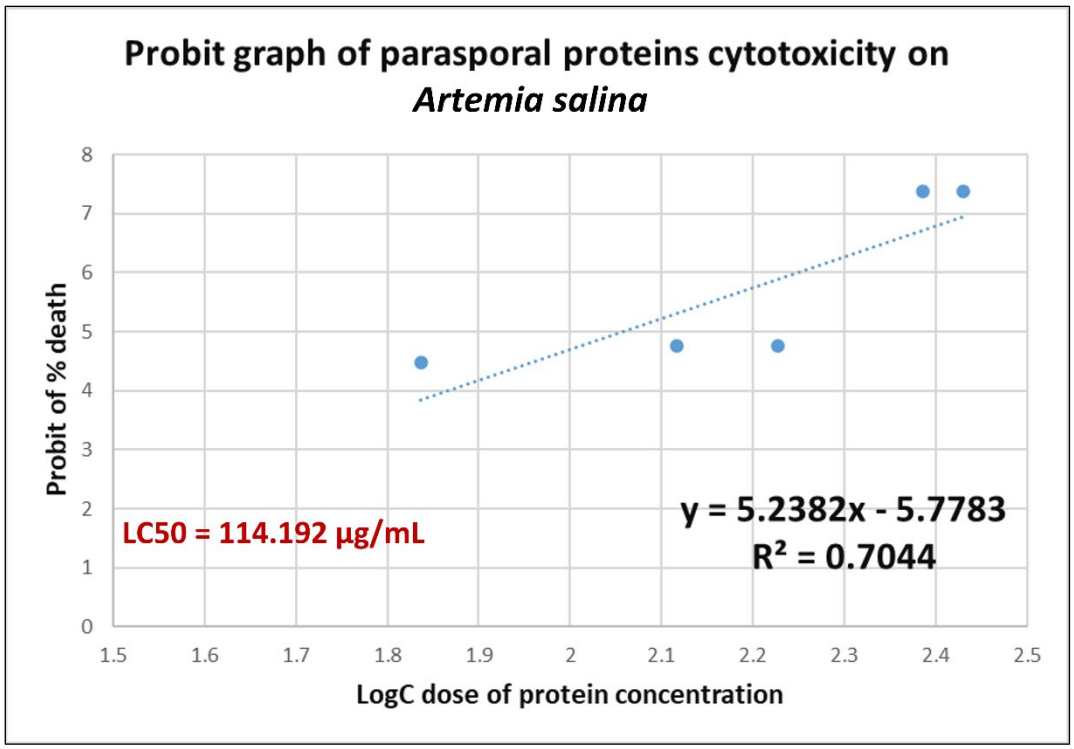



**Figure 5.3.5:** Probit graph of *in vivo* cytotoxicity of parasporal protein of indigenous Bt strains.

### 5.3.6 Hemolysis assays

The solubilized and untreated crude parasporal protein samples (Table 5.3.2) obtained from all five indigenous Bt strains were found to be nonhemolytic on sheep blood agar after being incubated at 37 °C for 18 hours (Figure 5.3.6).

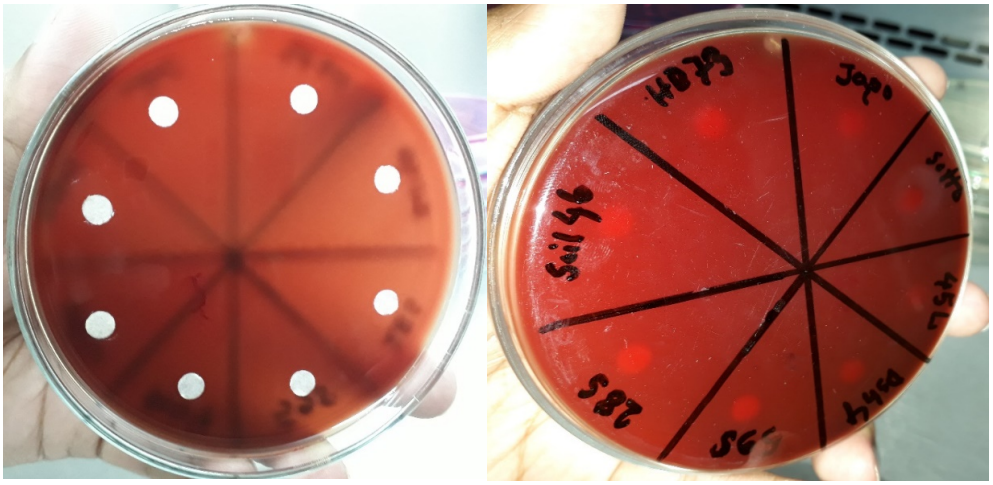



**Figure 5.3.6:** The hemolytic activity test conducted on solubilized parasporal proteins of indigenous Bt strains. BD59S, 28S, Soil-46, Dsh4, and 45L did not reveal any hemolytic activity on sheep blood agar after 18 hours of incubation at 37 $^{o}$C.

### 5.3.7 Insecticidal activity test

The solubilized and untreated crude parasporal protein samples (Table 5.3.2) obtained from all five indigenous Bt strains were likewise discovered to be non-insecticidal to the larvae of the melon fly (*Bactrocera cucurbitae*) (Figure 5.3.7).

**Figure 5.3.7:** The insecticidal bioassay test of parasporal proteins from BD59S, 28S, Soil-46, Dsh4, and 45L showed no efficacy against fruit fly larvae after 48 hours of incubation.

### 5.3.8 Cytotoxic activity of untreated parasporal protein on HeLa cell line

Three distinct variants of solubilized parasporal inclusion protein samples, namely untreated, proteinase K treated, and trypsin treated, were individually examined from each of the 5 *B. thuringiensis* strains to assess their cytotoxic effects on HeLa and Vero cell lines. Figure 5.3.2 (A-B) displays the protein bands and Table 5.3.3 presents the concentrations. Out of the 5 *B. thuringiensis* strains, the crude sample of solubilized parasporal proteins of the BD59S did not show any cytotoxic effects on the HeLa cell line (Figure 5.3.8). On the other hand, the crude protein samples from Bt strain 28S, Soil-46, 45L, and Dsh4 showed weak cytotoxicity towards the HeLa cell lines.

**Table 5.3.3:** Protein concentrations employed in the cell line experiments

| Indigenous Bt strains | Protein extracted from 50 mL culture (mg/mL) | Protein (µg/mL) | Protein added in each well (µg) | The final concentration of protein in each well (µg/µL) |
|---|---|---|---|---|
| **45L** | 1.546 | 1546 | 38.65 | 0.773 |
| **Dsh4** | 2.07 | 2070 | 51.75 | 1.035 |
| **BD59S** | 1.302 | 1302 | 32.55 | 0.651 |
| **28S** | 1.746 | 1746 | 43.65 | 0.873 |
| **Soil-46** | 1.934 | 1934 | 48.35 | 0.967 |

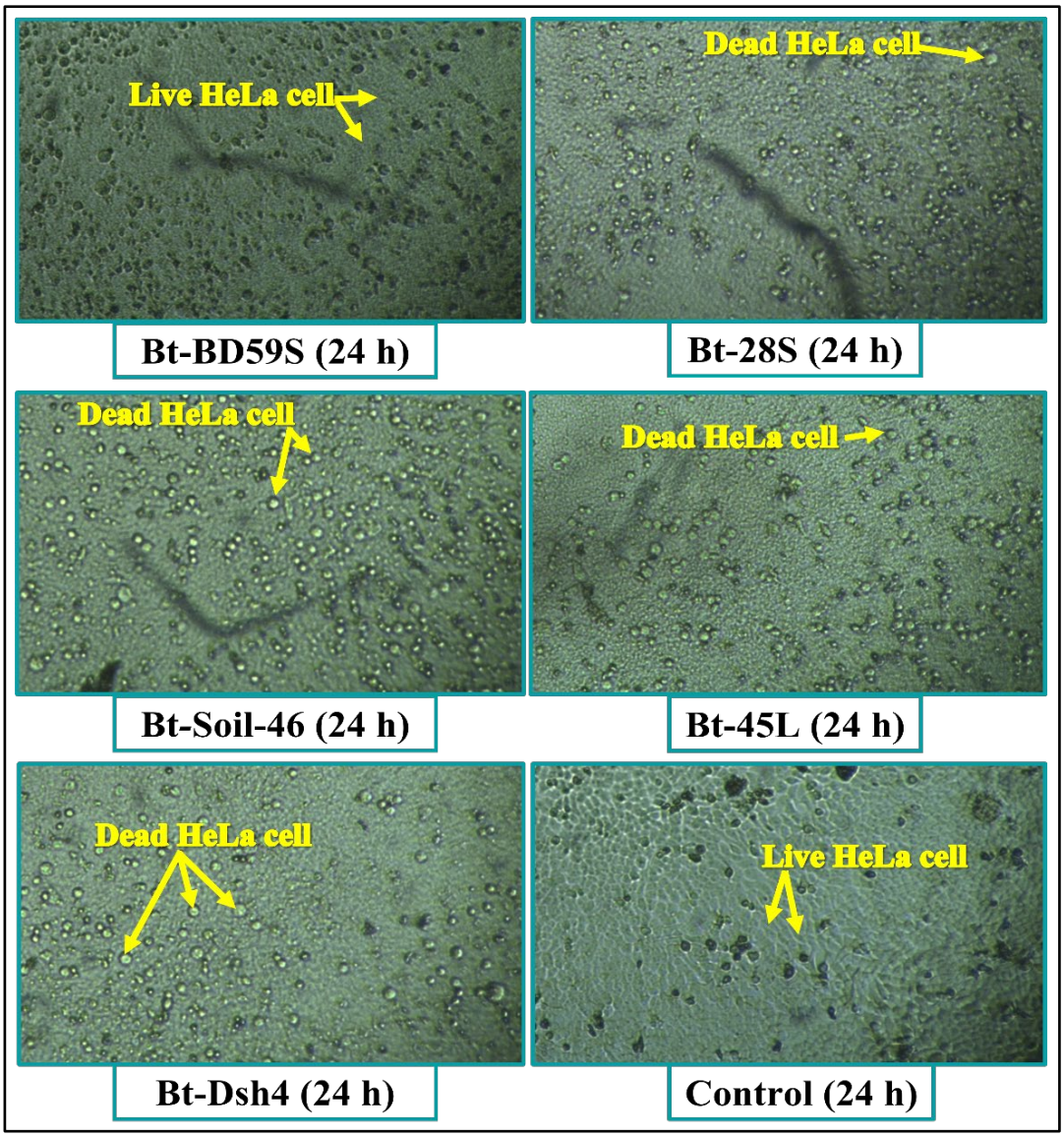



**Figure 5.3.8:** Cytotoxicity test of solubilized crude parasporal inclusion protein on HeLa cell line. BD59S did not show cytotoxicity, whereas 28S, Soil-46, 45L, and Dsh4 strains showed weak cytotoxicity on HeLa cell line after 24 hours of incubation.

### 5.3.9 Cytotoxic activity of trypsin digested parasporal protein on HeLa cell line

Among five *B. thuringiensis* strains, the trypsin-digested parasporal inclusion proteins of the BD59S exhibited significant cytotoxicity towards the HeLa cancer cell line (Figure 5.3.9). On the other hand, 28S, Soil-46, 45L, and Dsh4 showed only moderate cytotoxicity towards the HeLa cancer cell lines. BD59S exhibited more cytotoxicity than other strains, even though the crude protein of BD59S at the same concentration did not display any cytotoxic effects.

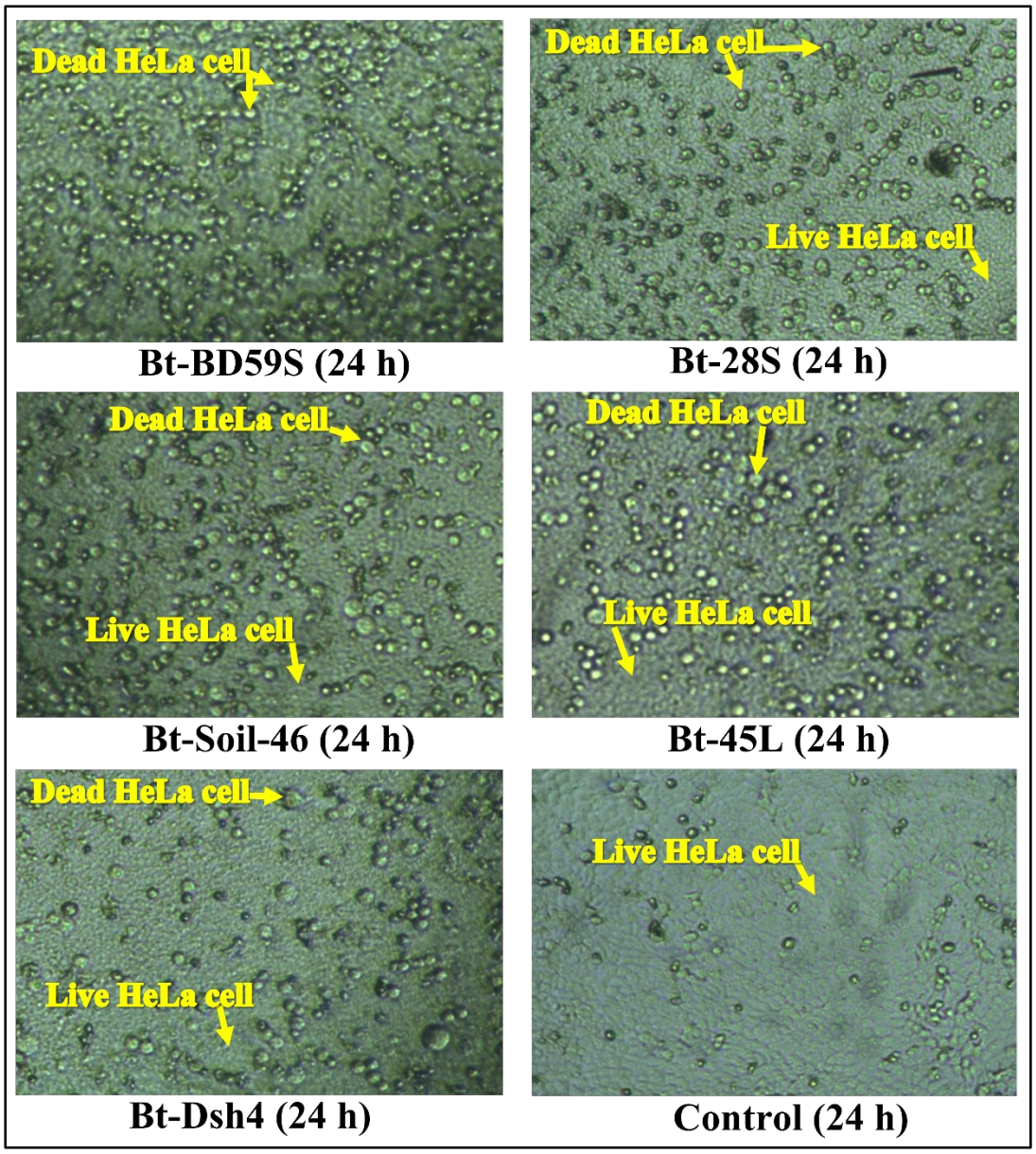



**Figure 5.3.9:** Cytotoxicity test of trypsin digested parasporal protein of five indigenous Bt strains on HeLa cell lines. BD59S showed moderate, whereas 28S, Soil-46, Dsh4, and 45L showed weak cytotoxicity on the HeLa cell line after 24 hours of incubation.

### 5.3.10 Cytotoxic activity of proteinase K digested parasporal protein on HeLa cell line

Proteinase K digested ~46 kDa parasporal proteins of each of 5 indigenous *B. thuringiensis* strains showed strong cytotoxic effects on the HeLa cell lines (Figure 5.3.10) than trypsin digested and crude protein. The presence of round-shaped deceased HeLa carcinoma cells suggests that they probably died due to necrosis.

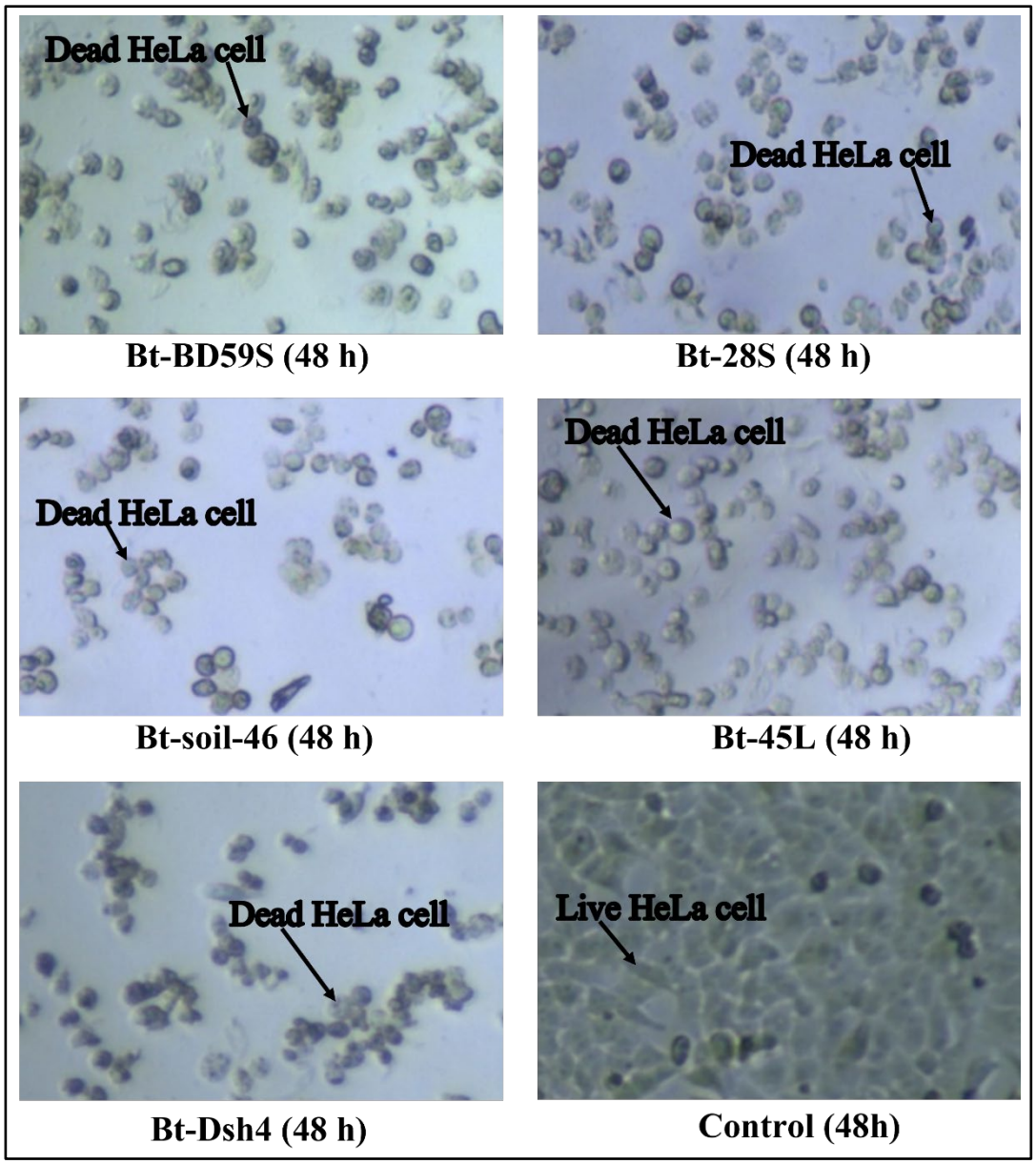



**Figure 5.3.10:** Cytotoxicity test of proteinase K digested parasporal protein of five indigenous Bt strains on HeLa cell lines. Proteinase K digested parasporal protein of BD59S, 28S, Soil-46, Dsh4, and 45L showed potent cytotoxicity against the HeLa cell line after 48 hours of incubation.

### 5.3.11 Cytotoxicity test of proteinase K treated parasporal proteins on Vero cell line

Following the digestion of parasporal proteins by proteinase K, the cytotoxicity of 5 native *B. thuringiensis* strains was found to be particularly effective against HeLa cell lines. Consequently, the cytotoxic effects of these strains on non-cancerous Vero cell lines were also investigated. Proteinase K digested parasporal proteins of 5 indigenous *B. thuringiensis* strains did not show any cytotoxic effects on the Vero cell lines (Figure 5.3.11).

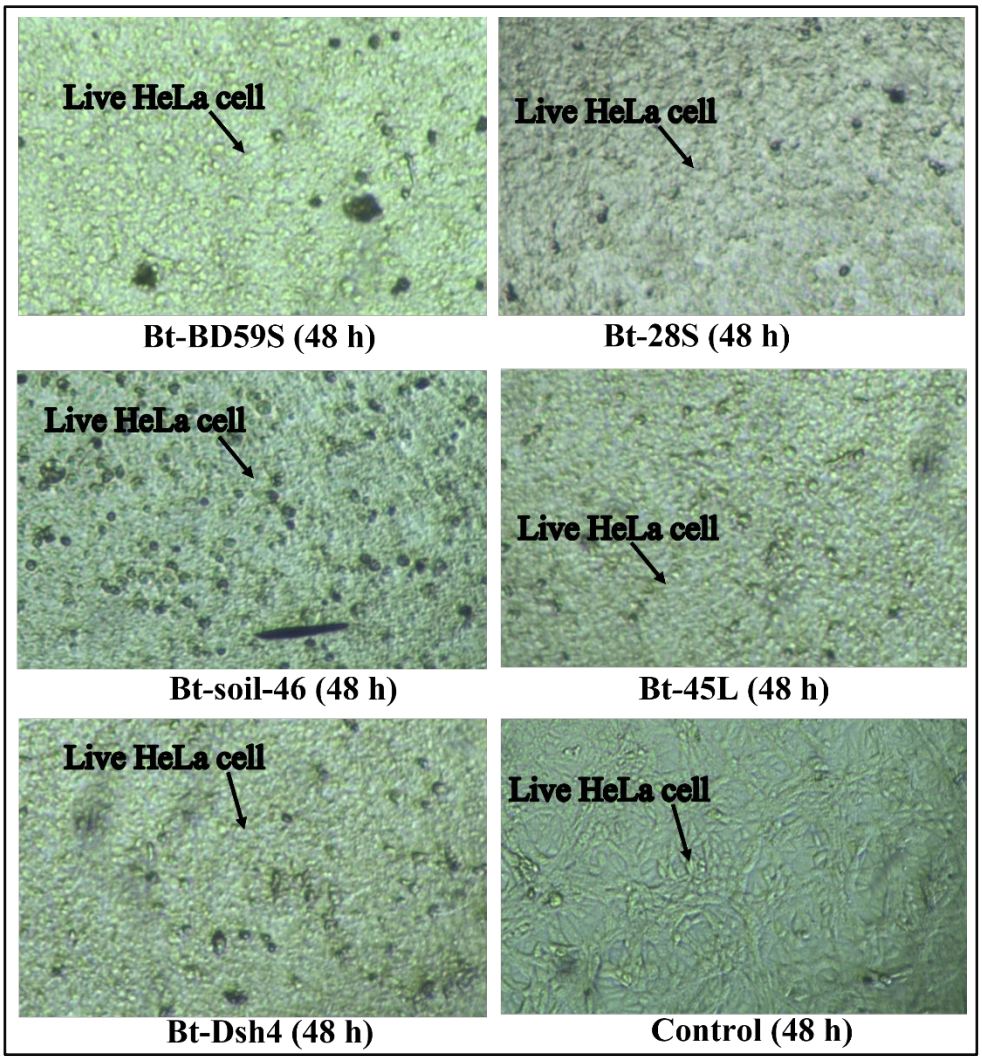



**Figure 5.3.11:** Cytotoxicity test of proteinase K digested parasporal protein of five indigenous Bt strains on Vero cell line. BD59S, 28S, Soil-46, Dsh4, and 45L did not show cytotoxicity against the Vero cell line after 48 hours of incubation.

### 5.3.12 Parasporal inclusion protein observation under the light phase microscopy

A study was conducted to observe the parasporal inclusions using a light-phase microscope. To achieve this objective, sonicated spore-crystal mixtures of 5 native *B. thuringiensis* strains were stained with 0.2% Coomassie blue, then were analyzed using light field of a phase-contrast microscope. As a result, somewhat round or oval or atypical shaped parasporal inclusion bodies were detected in each of the strains. The sole distinguishing factor between *B. thuringiensis* and other bacteria in the *B. cereus* group is in its ability to produce parasporal inclusion proteins. These findings, the presence of parasporal inclusions, verified that each of these *Bacillus* strains belonged to *B. thuringiensis* strains (Figure 5.3.12).

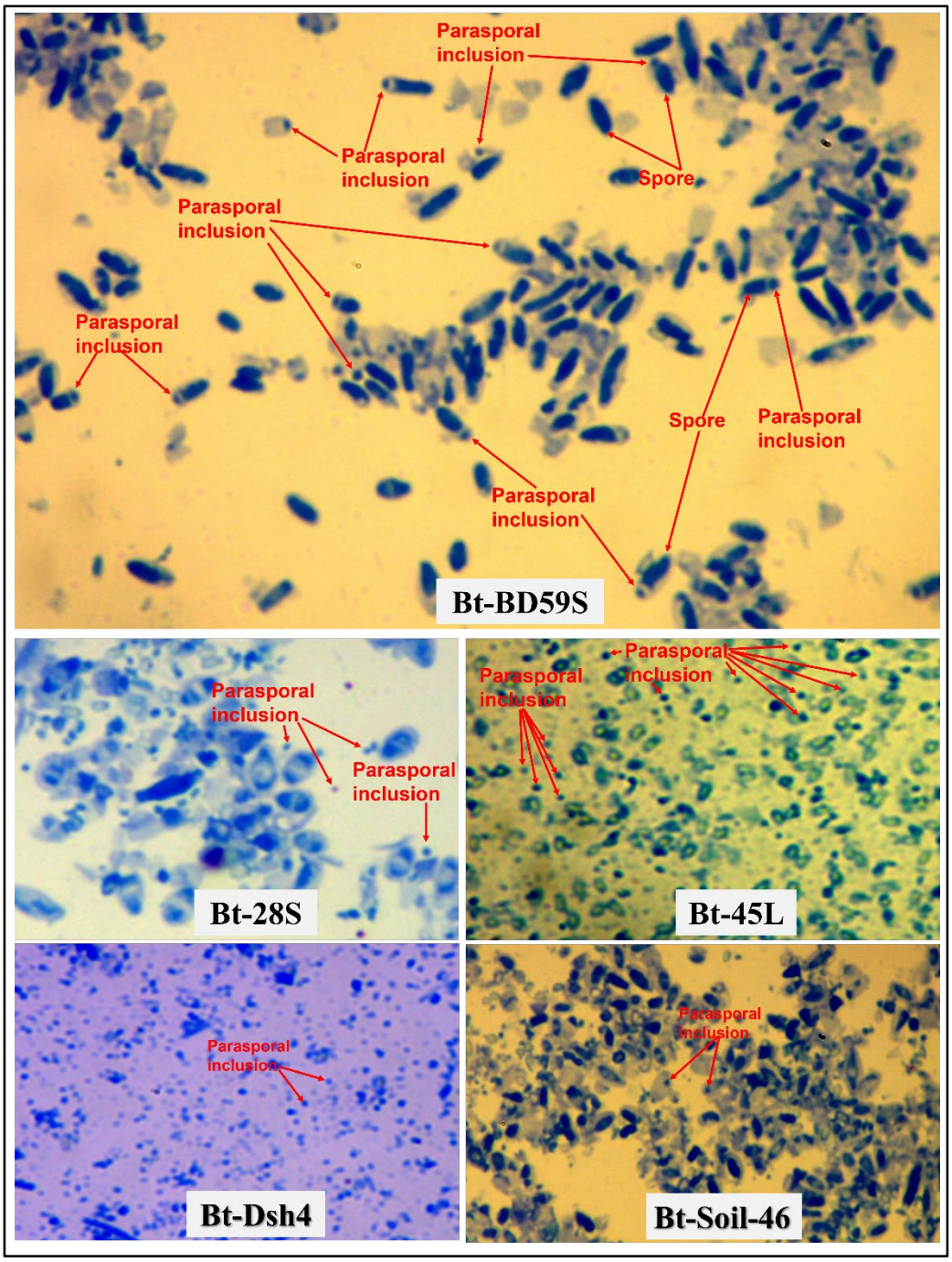



**Figure 5.3.12:** Coomassie blue (0.2%) stained parasporal crystalline inclusion protein of indigenous Bt strains BD59S, Soil-46, 28S, 45L, and Dsh4.

### 5.3.13 Parasporal inclusion protein observation from scanning electron microscope (SEM)

The SEM observation (Figure 5.3.13) provided clear evidence that each of the five indigenous *B. thuringiensis* strains (BD59S, Soil-46, 28S, 45L, and Dsh4) possesses parasporal inclusions. From SEM image, parasporal inclusions of all the strains were identified as round shaped with smooth surface. In addition, it was also found that there

was a variation in the size of the inclusion bodies, and the maximum diameter of the inclusion was measured as approximately 0.8 µm. The round-shaped parasporal inclusion closely resembled with the parasporal inclusion bodies from *B. thuringiensis* strain CTC, which is surface layer protein (Zhu & Yu, 2008).

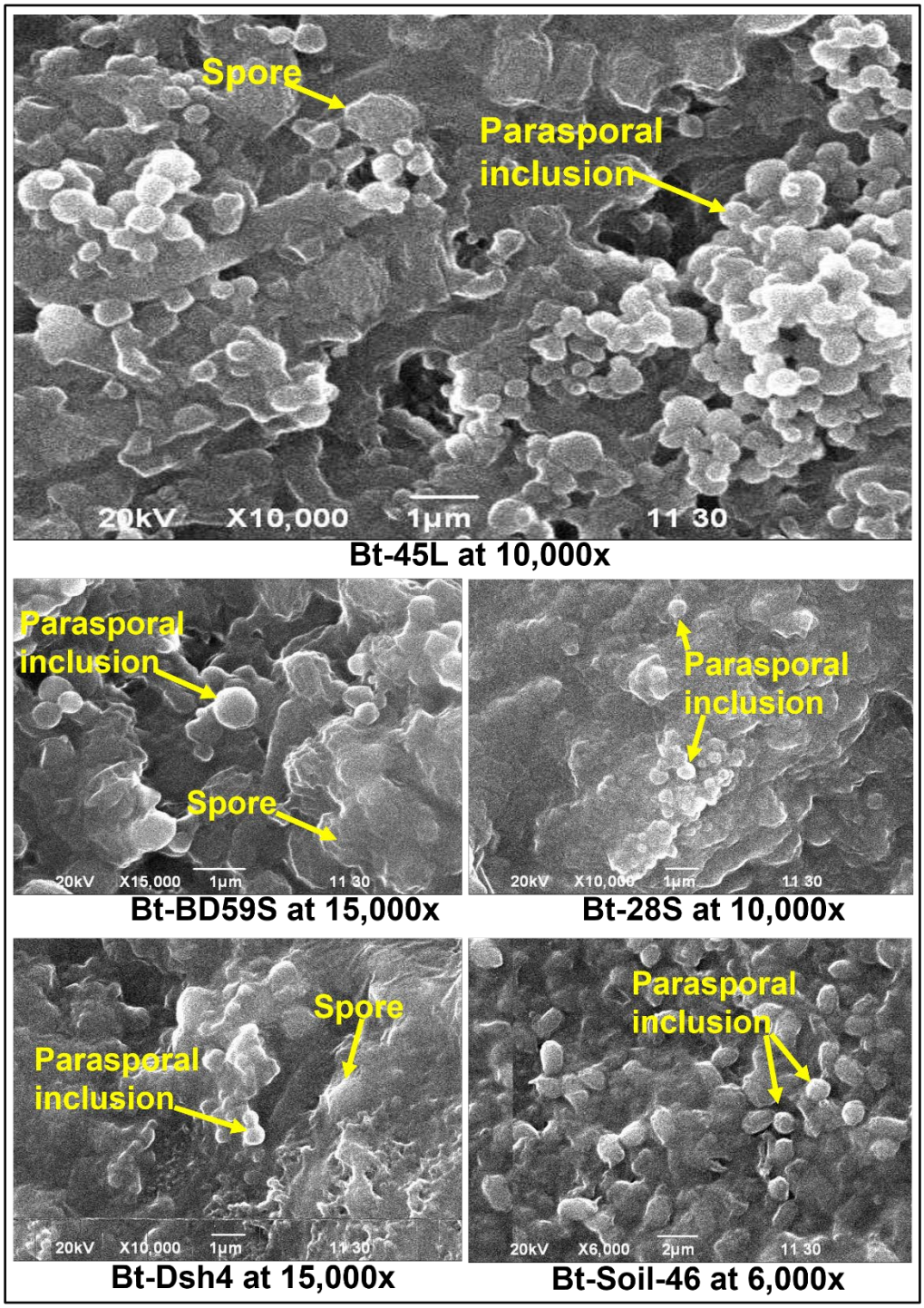



**Figure 5.3.13:** SEM observation of parasporal crystalline inclusions of indigenous Bt strains BD59S, Soil-46, 28S, 45L, and Dsh4. Parasporal inclusions of most of the indigenous Bt strains were found as round-shaped.

### 5.3.14 Colony characteristics of parasporal protein synthesizing *B. thuringiensis* strains

The strains were inoculated onto LB agar for 18 hours at 30 ºC to observe their colony morphology. The colonies were white to off-white, opaque, slightly raised elevation with a regular outline Figure 5.3.14. A summary of colony characteristics is shown in Table 5.3.4.

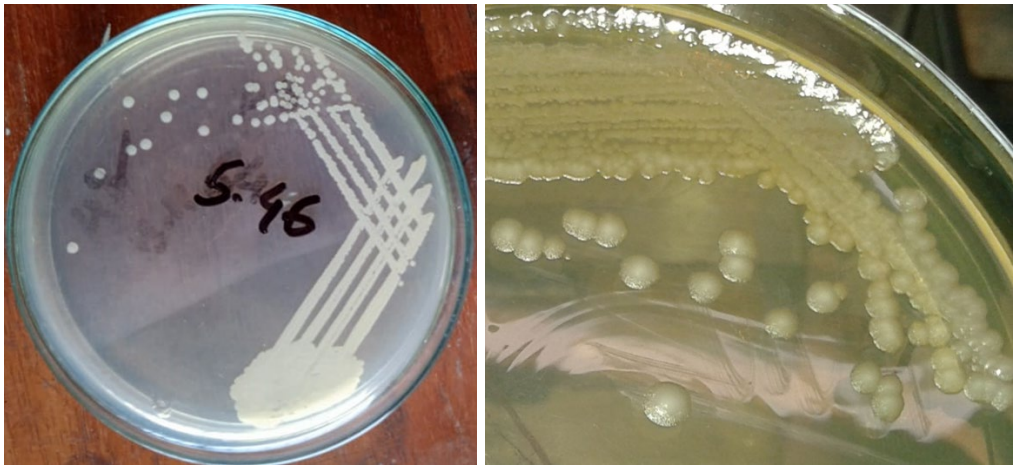



**Figure 5.3.14:** Colony characteristics of *B. thuringiensis* strain Soil-46 on LB agar plate

**Table 5.3.4:** Colony characteristics of *Bacillus thuringiensis* strains

| ***Parameters*** | ***Colony characteristics*** |
|---|---|
| *Size* | Medium to large |
| *Shape* | Round |
| *Elevation* | Raised |
| *Margin* | Wooly |
| *Surface* | Rough-smooth |
| *Color* | Off white |
| *Opacity* | Opaque |
| *Texture* | Non-gummy |

### 5.3.15 Biochemical characterization of parasporal protein synthesizing *B. thuringiensis* strains

Selective biochemical characterization test results of parasporal protein synthesizing indigenous 5 Bt strains, BD59S, 28S, Soil-46, 45L, Dsh4, and 3 reference Bt strains,

HD-73, sotto, japonensis were studied. A summary of biochemical test results is shown in Table 5.3.5.

**Table 5.3.5:** Biochemical characterization test results of parasporal protein synthesizing indigenous and reference *Bt* strains

| Strains / Biochemical test | HD-73 | sotto | japonensis | BD59S | 28S | Soil-46 | 45L | Dsh4 |
|---|---|---|---|---|---|---|---|---|
| **Gram staining** | + | + | + | + | + | + | + | + |
| **MR test** | - | - | - | - | - | - | - | - |
| **VP test** | + | + | + | + | + | + | + | + |
| **Indole test** | + | + | + | + | + | + | + | + |
| **Starch hydrolysis test** | + | + | + | + | + | + | + | + |
| **7% NaCl test** | + | + | + | + | + | + | + | + |
| **Gelatinase test** | + | + | + | + | + | + | + | + |

The main objective of our study in this chapter was to search for parasporin class 2, 4, and 5 or similar Cry proteins from indigenous non-insecticidal hemolytic Bt strains. We could obtain some prominent parasporal proteins with molecular mass of 103, 101, 86, and 76 kDa from 5 strains, namely, BD59S, Soil-46, 28S, 45L, and Dsh4 as observed on the SDS-PAGE. Upon digestion with proteinase K, or trypsin, or some other protease enzymes, these proteins get fragmented into relatively smaller active toxin protein. When we digested the proteins under study, with trypsin and proteinase K, they behaved like Cry proteins. All these parasporal proteins obtained in this study showed anticancer activity against HeLa cancer cells like parasporin and simultaneously showed non-cytotoxicity against non-cancer Vero cells. In addition, these proteins are nonhemolytic, and non-insecticidal as expected. Therefore, we initially considered all these parasporal proteins as parasporins. However, the shapes of parasporal inclusions under SEM match with the shapes of surface layer parasporal inclusion proteins. Therefore, whole genome sequencing of the strains to identify and characterize the proteins was the next logical step forward.

# CHAPTER 6:

# Whole genome sequencing of indigenous *Bacillus thuringiensis* strains and identification of their parasporal protein sequence

# CHAPTER 6: Whole genome sequencing of indigenous *Bacillus thuringiensis* strains and identification of their parasporal protein sequence

## 6.1 Introduction

Bacterial whole-genome sequencing (WGS) by next generation sequencing (NGS) is a rapid, cost-effective, and highly efficient important technology for knowing the entire bacterial genome information, genotyping and genetic features including virulence, toxins, bacteriocins, antibiotic-resistance genes, any gene of interest, and their metabolic and biological subsystems. In our study, no definite parasporin or other cry gene sequences were identified by PCR, although the parasporal proteins with parasporin-like anticancer activity were found from 5 indigenous *B. thuringiensis* strains. To determine their genotype and protein sequence of the parasporal inclusions and its corresponding genes, whole genome sequencing of four indigenous *B. thuringiensis* strains (BD59S, 28S, Soil 46, and 45L) were conducted using the MiSeq System of Illumina NGS technology. This chapter contains the resulting WGS data that were assembled, annotated, and analyzed for the identification and characterization of parasporal proteins of indigenous *B. thuringiensis* strains.

## 6.2 Materials and Methods

### 6.2.1 Materials

#### 6.2.1.1 Chemicals

The chemicals that were used for this research are detailed in **Appendix III.**

#### 6.2.1.2 Types of equipment

**Appendix IV** includes information on the manufacturers and types of equipment utilized in this research, which is detailed in the methods section.

#### 6.2.1.3 Media

The media that were used for this research are detailed in **Appendix I.**

#### 6.2.1.4 Reagents and solutions

The methods section included descriptions of the reagents and solutions used in the experiments, whereas Appendix II detailed the composition and production methods of these substances.

#### 6.2.1.5 Bacterial strains

Four (4) different indigenous *B. thuringiensis* strains **BD59S, 28S, Soil 46**, and **45L** were studied.

### 6.2.2 Methods

For bacterial whole-genome sequencing (WGS) Illumina MiSeq technology was preferred to identify the SDS-PAGE observed parasporal anticancer protein synthesizing genes and protein sequence as well as to know the genome information of 4 indigenous *B. thuringiensis* strains 45L, BD59S, 28S, and Soil 46.

#### 6.2.2.1 Genomic DNA extraction

The bacterial strains were cultured overnight at 37 °C in 100 mL of LB broth to prepare the samples for sequencing. GeneJET Genomic DNA Purification Kit, manufactured

by Thermo Scientific was used to isolate the DNA. Centrifugation was used for 10 minutes at 5000 × g to collect about 2×109 bacterial cells from each strain in four separate 1.5 microcentrifuge tubes. The culture supernatants were thrown out and the pellets were mixed with 180 µL of Gram-positive bacteria lysis solution (20 mM Tris-HCl, pH 8.0, 2 mM EDTA, 1.2 percent Triton X-100, and 20 mg/mL lysozyme). Following that, it was left to incubate at 37 °C for 30 minutes. To create a consistent suspension, 200 µL of Lysis Solution and 20 µL of Proteinase K were added and stirred completely using either pipetting or vortexing. The samples were then incubated in a shaking water bath at 56 °C for about 30 minutes, or until the cells were fully lysed. After that, 20 µL of RNase-A solution was combined by vortexing, and the resulting mixture was left to incubate at room temperature for 10 minutes. Next, combine 400 µL of 50% ethanol by pipetting them together. A GeneJET genomic DNA purification column was then inserted into a collecting tube containing the lysate. Spin the column at 6,000 × g. After draining the solution from its collecting tube, it was discarded. The genomic DNA purification column from GeneJET was transferred to a fresh 2 mL collection tube. After discarding the flow-through, the purification column was reinserted into the collecting tube. Using 500 µL of Wash Buffer II, the GeneJET genomic DNA purification column was filled. The next step was a three-minute centrifugation run at maximum speed (≥12000 × g). The flow-through solution was removed from the collecting tube and discarded and the column was transferred to a new 1.5 mL microcentrifuge tube. To extract the genomic DNA, 200 µL of elution buffer was added to the middle of the membrane of the GeneJET genomic DNA purification column. Following that, it was centrifuged for 1 minute at 8000 × g after being incubated for 2 minutes at room temperature. The column used for purification was thrown away. Before preparing the genome sequencing library, the DNA was measured using NanoDrop One from Thermo Fisher Scientific in the USA and kept at -20 °C.

#### 6.2.2.2 Bacterial genome sequencing

Using the Nextera XT DNA library preparation kit (Illumina), a 300-bp paired-end library was constructed from the genomic DNA. Then, they were normalized using bead-based technology, following the manufacturer's methodology. The DNA of each strain was then sequenced at the Genome Research Institute of North South University

(NSU), Bangladesh, using the MiSeq Reagent Kit v3 (600-cycle) by the Illumina MiSeq technology (Illumina, San Diego, CA, USA). The trimming of the Illumina adaptor was done using BaseSpace v3.

#### 6.2.2.3 Quality assessment, trimming, and de novo genome assembly

Trimmomatic v0.38 (Bolger *et al.*, 2014) was used to do quality filtering of the reads in a sliding window mode (window size: 4:15) to confirm a minimum length of 35 and a minimum Phred score of 33. FastQC v0.11.7 was then used to examine the quality-filtered paired-end readings (Andrews, 2010). Following that, SPAdes v3.13 (Bankevich *et al.*, 2012) was used to assemble paired-end reads, and the "careful" mode was set by default for k-mer lengths. The resulting contigs were then mapped and moved using the move contigs command of mauve v2.4.0 (Darling *et al.*, 2004) with default parameters to get the single circular genome from the scaffolded multi-fasta file of each genome. Each genome was then organized by using the microbial genome blasted best-matched reference genome of the *B. thuringiensis* strain of the NCBI (National Center for Biotechnology Information) database.

#### 6.2.2.4 Genome annotations and analyses

Annotations were performed using three different annotation webservers, the RAST (Rapid Annotation using Subsystem Technology), the BV-BRC (Bacterial and Viral Bioinformatics Resource Center), and the NCBI PGAP (Prokaryotic Genome Annotation Pipeline). The RAST server was used to carry out the functional annotation of the genetic features of the organized contigs of the genomes (Aziz *et al.*, 2008). In this case, errors were automatically fixed with Classic RAST v2.0 and FIGfam v70 were used. Besides, metabolic models were developed and gaps were backfilled. In RAST annotation both replication and the debug were turned off. Besides, genome annotations were also performed from the BV-BRC web server (Olson *et al.*, 2023) and subsystems, pathways, phylogeny, genome browser, circular genome viewer, comparative genetic features, virulence factor, transporter, drug target, antibiotic resistance, etc. were predicted and studied using this web server. Besides, comprehensive genome analysis was also performed from BV-BRC. A prokaryotic genome annotation pipeline (PGAP) from NCBI was used for final annotations (Tatusova *et al.*, 2016). Artemis was used to view the annotated characteristics of

plasmids and genomes (Rutherford *et al.*, 2000). Using BV-BRC, Gview (Petkau *et al.*, 2010), and DNAPlotter (Carver *et al.*, 2009), circular graphical maps of genomes and plasmids were created to show the areas and features.

#### 6.2.2.5 Comparative genome analysis

The NCBI microbial genome BLAST (Basic Local Alignment Search Tool), and BV-BRC similar genome finder were used to identify the native Bt strains and their closely related species. In addition, a genome comparison with closely related species was determined using BRIG (Blast Ring Image Generator) version 0.95 (Alikhan *et al.*, 2011). BRIG was used to create the circular comparative genomic map using NCBI local blast-2.8.1+ with the standard default parameters. Multiple genome alignment was performed by progressive mauve. Furthermore, using the genome alignment guide tree of the progressive mauve genome alignment, phylogenetic relatedness between closely related species of B. cereus and *B. thuringiensis* was identified (Darling *et al.*, 2010), and FigTree v1.4.0 was used to display the phylogenetic genomic tree (Rambaut, 2010).

#### 6.2.2.6 Identification of parasporal proteins of indigenous Bt strains

All parasporin and *cry* gene sequence were collected from NCBI database according to the list of parasporins (https://www.fitc.pref.fukuoka.jp/parasporin/list.html), and list of Cry (old website: http://www.lifesci.susx.ac.uk/home/Neil_Crickmore/Bt/ new website: https://www.bpprc-db.org/home/). Then a multi-fasta reference file was created by merging all of *ps* and *cry* gene sequences. This reference multi-fasta file was then aligned with the NGS paired-end fastq short-read sequence file of each genome using two alignment tools, BWA (Li & Durbin, 2009), and the Bowtie 2 (Langmead & Salzberg, 2012), in ubuntu operating system. Then obtained SAM alignment file was converted to BAM and sorted BAM from samtools v1.9 (Li *et al.*, 2009). Resulted alignment was then observed using sequence assembly visualization tool Tabet (Milne *et al.*, 2009).

From the NCBI-PGAP, BV-BRC, and RAST annotated genome sequencing data of indigenous *B. thuringiensis* strains, protein sequences those molecular weights ranging from 60 kDa to 125 kDa were calculated manually from the protein molecular weight calculator (https://www.bioinformatics.org/sms/prot_mw.html) and then checked by

NCBI protein BLAST (Johnson *et al.*, 2008), UniProt (Universal Protein Resource) BLAST (Pundir *et al.*, 2016), and InterProScan (Mulder *et al.*, 2007) web database for their key identification.

#### 6.2.2.7 Multiple sequence alignment, percent identity, and phylogenetic tree construction

Multiple sequence alignment (MSA) was performed by Clustal Omega (Sievers & Higgins, 2018), and MAFFT (Multiple Alignment using Fast Fourier Transform) version-7 (Katoh, Rozewicki, & Yamada, 2018) online tools. Resulted fasta alignment files were used to construct the phylogenetic trees. Initially, the phylogenetic trees were constructed by both Clustal Omega and MAFFT webserver. Finally, phylogenetic trees were constructed by the MAFFT web server using the neighbor-joining (NJ) method with 1000 bootstraps. The resulting newick formatted tree files were downloaded viewed and designed by the FigTree v1.4.0. Besides, both alignment and percent identity of possible parasporal protein sequences of indigenous Bt strains with parasporal inclusion forming S-layer protein, anticancer parasporins, and insecticidal Cry proteins were calculated and viewed by the MView (Brown *et al.*, 1998) of EBI (European Bioinformatics Institute) webserver. MView resulted from alignment files with percent coverage and identity that were downloaded by printing them as pdf. Besides, percent identity was also determined by the Clustal Omega result from the alignment file, where the percent identity matrix was calculated by Clustal2.1 from the Clustal Omega alignment result summary.

#### 6.2.2.8 Identification of genetic organization of features in the genome

The genetic organization of desired features in the genome was observed from the BV-BRC webserver. The desired gene and protein name or id was found from the BV-BRC annotated or NCBI PGAP annotated desired bacterial strains genome sequence data and then searched in the BV-BRC webserver. Then the desired feature was selected and detail information was obtained from the switch to feature view. Then the feature was searched by the compare region viewer menu. A result page with the comparative genetic organization of the desired feature in the desired genome and the same gene or proteins in the other genome originated. The desired gene or feature position was observed within the region from 5000 to 100000 bp, and finally, the best-viewed result was exported as a svg file.

#### 6.2.2.9 Comparative motif analysis

Motifs are small patterns of conserved sequence that are linked to certain protein activities. The conserved motifs of the parasporal inclusion forming proteins were investigated using the online MEME (Multiple Expectation maximizations for Motif Elicitation) suite 5.5.5 (Bailey *et al.*, 2009). A multi-fasta of desired protein sequences was prepared by ordering them according to their position in the phylogenetic tree from top to bottom. Then motif search for the batch protein sequences was run using the default settings of the classic motif discovery mode of the MEME suite, with a maximum number of 10 motifs for each protein sequence. The resulting HTML, XML, and Txt files were saved and observed. The comparative motif patterns were then visualized using the XML file from the bio sequence structure illustrator of the graphics menu of the TBtools v1.130 (Chen *et al.*, 2020). The resulting motif architecture was then saved as a jpg or png file and associated with the phylogenetic tree and analyzed.

#### 6.2.2.10 Comparative domain analysis

Protein domains are the fundamental building blocks of proteins and can fold, function, and develop on their own. Understanding protein domains is essential for classifying proteins, recognizing their biological functions, identifying their evolutionary mechanisms, and designing proteins. The NCBI-CDC (Conserved Domain Search) program (Marchler-Bauer *et al.*, 2015) was used to investigate the conserved domains of the parasporal inclusion-producing proteins. A multi-fasta of desired protein sequences was prepared by ordering them according to their position in the phylogenetic tree from top to bottom. The domain search was run using the default settings of the Pfam v34.0 database on the NCBI-CDC tool with a 0.01 expect value (E-value) threshold and a maximum number of hits was set for 500 for the batch protein sequences. Then the resulting full domain hits data was downloaded as an ASN text file. The comparative domains were then visualized using the downloaded Txt file and proteins multi-fasta file from the bio sequence structure illustrator of the graphics menu of the TBtools v1.130 (Chen *et al.*, 2020). The resulting domain architecture was then saved as a jpg or png file and associated with the phylogenetic tree and analyzed.

#### 6.2.2.11 Identification of individual protein's domain by SMART tool

The SMART (Simple Modular Architecture Research Tool) domain search (Letunic *et al.*, 2021; Ponting *et al.*, 1999; Schultz *et al.*, 1998), was used to identify the schematic diagram of the domain architecture and domain sequence of probable parasporal inclusion forming proteins of indigenous Bt strains. Each protein sequence was analyzed from SMART by using both normal and genomic modes separately.

#### 6.2.2.12 *In silico* tertiary and secondary protein structure prediction and visualization

Proteins 3D structure homology modeling was performed using five online webservers Robetta (Kim *et al.*, 2004), I-TASSER (Yang *et al.*, 2015), Phyre2 (Kelley *et al.*, 2015), SWISS-MODEL (Waterhouse *et al.*, 2018), and MODELLER (Eswar *et al.*, 2006) for proper identification and characterization of the SDS-PAGE observed parasporal protein molecules. Then each of the predicted 3D pdb structures of proteins was assessed and evaluated by analysis of multiple parameters from the web server of EBI-PDBsum (Laskowski *et al.*, 2018), SWISS-MODEL (Waterhouse *et al.*, 2018), and SAVES v6.0 (https://saves.mbi.ucla.edu/) for identification of the best-modeled protein. Proteins 3D pdb structures were visualized from PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC., n.d.). Besides, the secondary structure of each of the protein sequences was predicted and visualized from the modeled protein's 3d pdb structure using the EBI-PDBsum webserver.

## 6.3 Results

### 6.3.1 Quality assessment, trimming, and de novo genome assembly of indigenous *B. thuringiensis* strains

From the FastQC quality assessment report, it was found that sequence qualities were good; no adapter content, basic statistics, per-sequence quality scores, sequence length distribution, and per-base sequence quality of the NGS data were found to be suitable for genome assembly (Figure 6.3.1). A summary of the quality assessment of NGS data for the BD59S, 28S, Soil-46, and 45L strains is shown in Table 6.3.1. Besides, de novo genome assemblies were performed successfully by the SPAdes genome assembler, and a summary of the SPAdes v3.13 de novo genome assemblies of these strains are shown in Table 6.3.2.

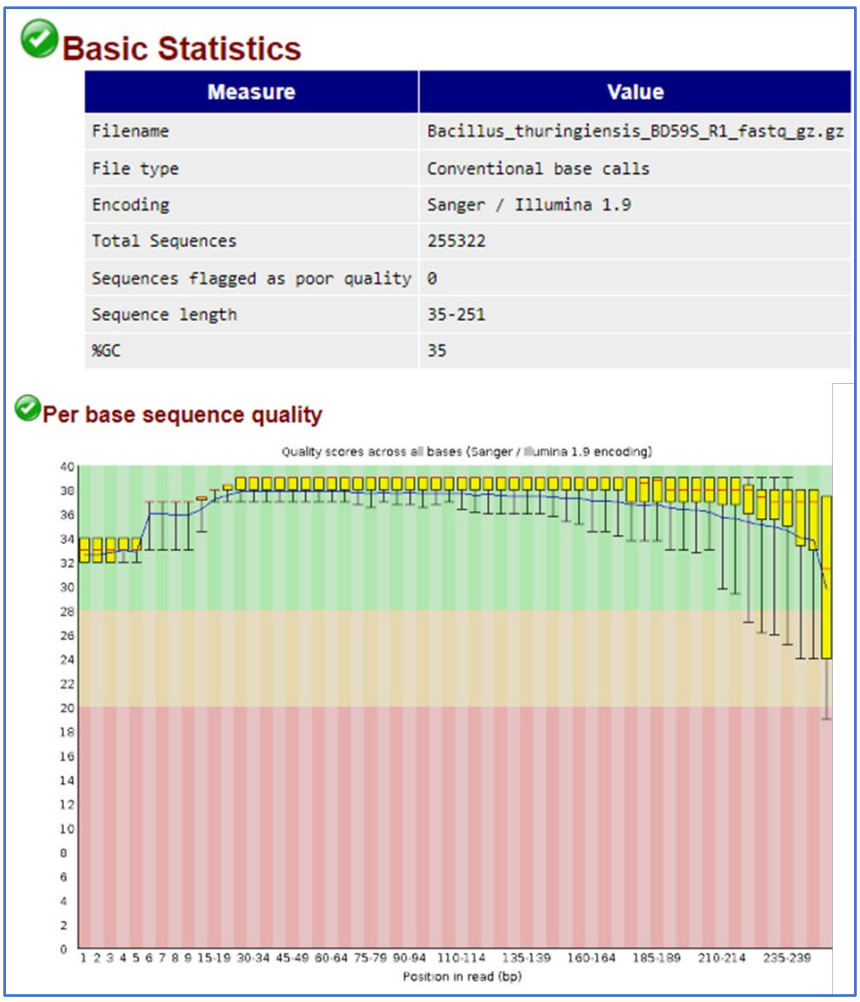



**Figure 6.3.1:** FastQC quality assessment result of BD59S after trimming and adapter removal.

**Table 6.3.1:** Summary of quality assessment of NGS data for BD59S, 28S, Soil-46, and 45L strains

| Name of Bt strain | Read pairs | Average read length | GC (%) |
|---|---|---|---|
| BD59S | 255322 | 35-251 | 35 |
| 45L | 836027 | 151 | 35 |
| Soil-46 | 1577228 | 20-74 | 37 |
| 28S | 2469334 | 74 | 37 |

**Table 6.3.2:** Summary of features of BD59S, 28S, Soil-46, and 45L whole genome sequence assemblies

| Name of Bt strain | Genome assembler (*de novo*) | Genome sequence length (bp) | GC (%) | Number of contigs | Longest contig length (bp) | Contigs N50 |
|---|---|---|---|---|---|---|
| BD59S | SPAdes 3.13 | 5,283,933 | 35.4 | 148 | 355048 | 144600 |
| 45L | SPAdes 3.13 | 5,375,636 | 35.1 | 82 | 580902 | 192714 |
| Soil-46 | SPAdes 3.13 | 5,278,071 | 35.3 | 67 | 564293 | 218747 |
| 28S | SPAdes 3.13 | 5,275,448 | 35.3 | 241 | 260043 | 97665 |

### 6.3.2 Genome annotations

A summary of genome annotation features obtained from NCBI PGAP v4.7 of the BD59S, 28S, Soil-46, and 45L strains is shown in Table 6.3.3. Table 6.3.4 shows a summary of the subsystem features of the full genome sequence of Bt strains BD59S, 28S, Soil-46, and 45L annotations. Figure 6.3.2 shows how the genomes cover different subsystems. Besides, the total annotated genes with their position and direction of translation are shown in the circular genome map (Figure 6.3.3).

**Table 6.3.3:** Summary of features of BD59S, 28S, Soil-46, and 45L whole genome sequence annotation

| Name of Genome | Total CDS | Sub systems | tRNA | rRNAs (5S, 16S, 23S) | ncRNAs | Pseudo Genes | GenBank Accession & Date |
|---|---|---|---|---|---|---|---|
| BD59S | 5810 | 356 | 98 | 12, 7, 6 | 5 | 260 | CP034684-86 (20-DEC-2018) |
| 45L | 5501 | 333 | 58 | 1, 4, 4 | 5 | 142 | JAQOOH0000 00000 (01-FEB-2023) |
| Soil-46 | 5436 | 338 | 30 | 1, 2, 3 | 5 | 122 | - |
| 28S | 5513 | 337 | 30 | 1, 2, 3 | 5 | 161 | - |

**Table 6.3.4:** Subsystem-wise features of BD59S, 28S, Soil-46, and 45L whole genome sequence annotation

| | *Bt*-BD59s | *Bt*-soil-46 | *Bt*-28s | *Bt*-45L |
|---|---|---|---|---|
| *Cofactors, Vitamins, Prosthetic Groups, Pigments* | 211 | 155 | 155 | 158 |
| *Cell Wall and Capsule* | 83 | 87 | 88 | 70 |
| *Virulence, Disease, and Defense* | 64 | 60 | 60 | 61 |
| *Potassium metabolism* | 10 | 10 | 10 | 10 |
| *Miscellaneous* | 29 | 28 | 30 | 26 |
| *Phages, Prophages, Transposable elements, Plasmids* | 20 | 20 | 20 | 8 |
| *Membrane Transport* | 71 | 41 | 46 | 39 |
| *Iron acquisition and metabolism* | 43 | 39 | 39 | 57 |
| *RNA Metabolism* | 57 | 55 | 56 | 60 |
| *Nucleosides and Nucleotides* | 122 | 117 | 119 | 112 |
| *Protein Metabolism* | 180 | 160 | 149 | 160 |
| *Cell Division and Cell Cycle* | 6 | 6 | 6 | 27 |
| *Motility and Chemotaxis* | 8 | 8 | 8 | 8 |
| *Regulation and Cell signaling* | 52 | 33 | 33 | 30 |
| *Secondary Metabolism* | 9 | 8 | 8 | 8 |
| *DNA Metabolism* | 69 | 69 | 69 | 79 |
| *Fatty Acids, Lipids, and Isoprenoids* | 78 | 64 | 64 | 68 |
| *Nitrogen Metabolism* | 7 | 7 | 7 | 16 |
| *Dormancy and Sporulation* | 111 | 110 | 114 | 92 |
| *Respiration* | 76 | 79 | 79 | 77 |
| *Stress Response* | 38 | 39 | 39 | 39 |
| *Metabolism of Aromatic Compounds* | 10 | 10 | 10 | 12 |
| *Amino Acids and Derivatives* | 417 | 361 | 371 | 348 |
| *Sulfur Metabolism* | 7 | 6 | 6 | 6 |
| *Phosphorus Metabolism* | 42 | 21 | 21 | 21 |
| *Carbohydrates* | 285 | 247 | 247 | 249 |

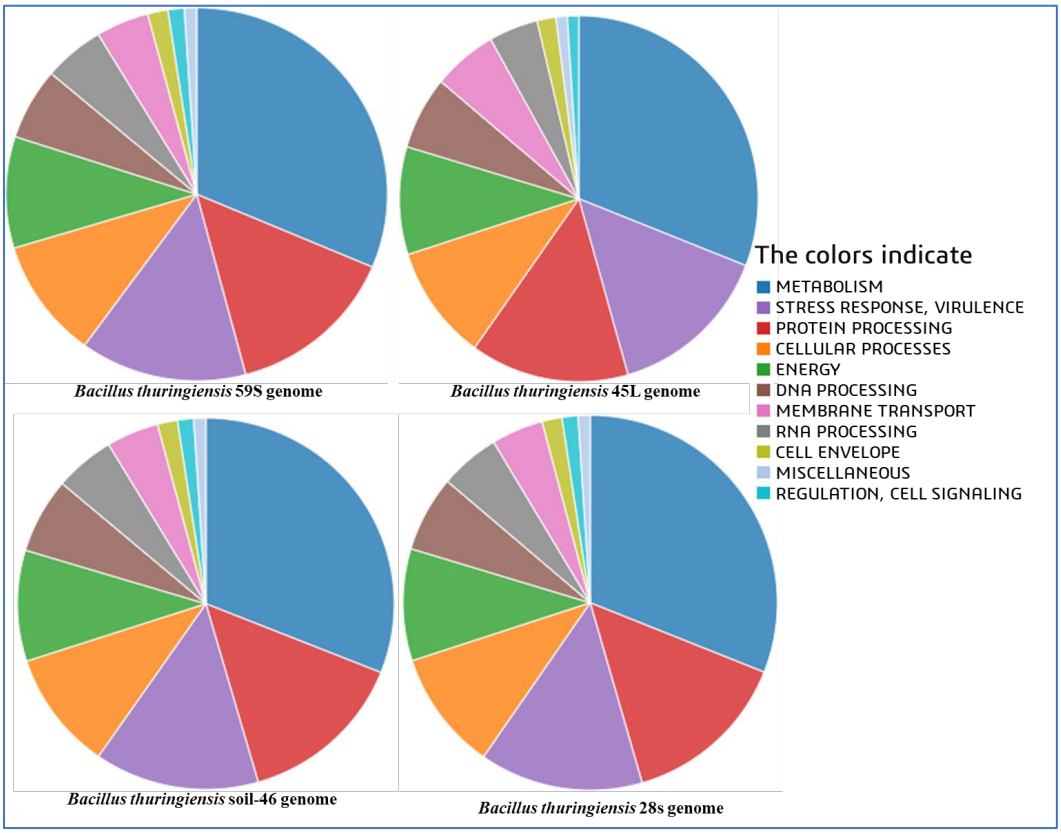



**Figure 6.3.2:** Pie charts showing coverage of different subsystems of the BD59S, 28S, Soil-46, and 45L genome annotation.

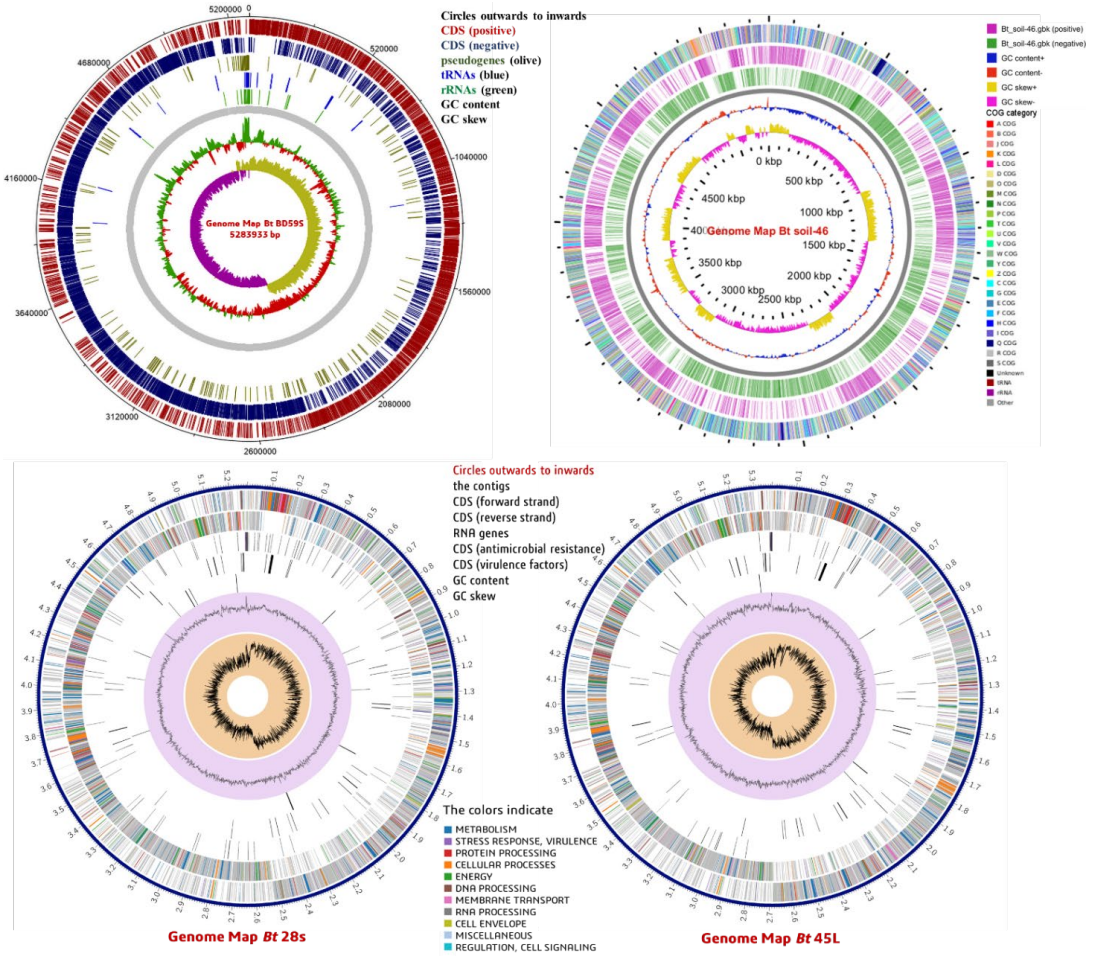



**Figure 6.3.3:** Circular genome maps with annotations of Bt strains BD59S, 28S, Soil-46, and 45L. (Generated by DNA plotter software, Gview, and the BV-BRC web server. Circles indicate, from inside to outwards: GC skew; GC content; rRNAs; tRNAs; pseudogenes; CDS on the reverse strand; CDS on the forward strand).

## 6.3.3 Comparative genome studies of indigenous *B. thuringiensis* with closely related *B. cereus* group bacteria

### 6.3.3.1 BLAST investigation by BRIG

To determine the identity of four indigenous Bt strains (BD59S, Soil-46, 28S, and 45L), the NCBI microbial genome BLAST revealed that all of these strains share 98–97% similarity with the complete genome sequence of fourteen *B. thuringiensis*, two *B. anthracis,* and three *B. cereus* strains. Further, NCBI local blast-2.8.1+ was used by BRIG (Alikhan *et al.* 2011) to visually compare the genome sequence identity, their similarity, and dissimilarity of indigenous Bt genomes with those of 19 genomes to identify regions of similarity and differences. From the BRIG result (Figure 6.3.4), it was found that among 4 indigenous Bt strains, 3 (BD59S, Soil-46, and 28S) were more identical than 45L. Besides, all of the genomes compared here have mostly similar and

few dissimilar regions among them. From the circular graphical BRIG image (Figure 6.3.4), it was found that the genomes of indigenous 4 Bt strains had more similarities with the genomes of 6 Bt strains: CTC (CP013274), *konkukian* 97-27 (NC_005957), LM1212 (CP024771), BM-BT15426 (CP020723), HS18-1 (CP012099), and Coreansis-ST7 (CP016194). Besides these, the genomes of Bt strain XL6 (CP013000) had two large dissimilar regions with all of the Bt, Bc, and Ba genomes. Besides, *B. anthracis* and *B. cereus* also had close similarities with *B. thuringiensis* genomes, but 2 *B. anthracis* (CP015779, CP009902) and Bt HD682 (CP009720) had a close common region that differs them from other Bt and Bc genomes.

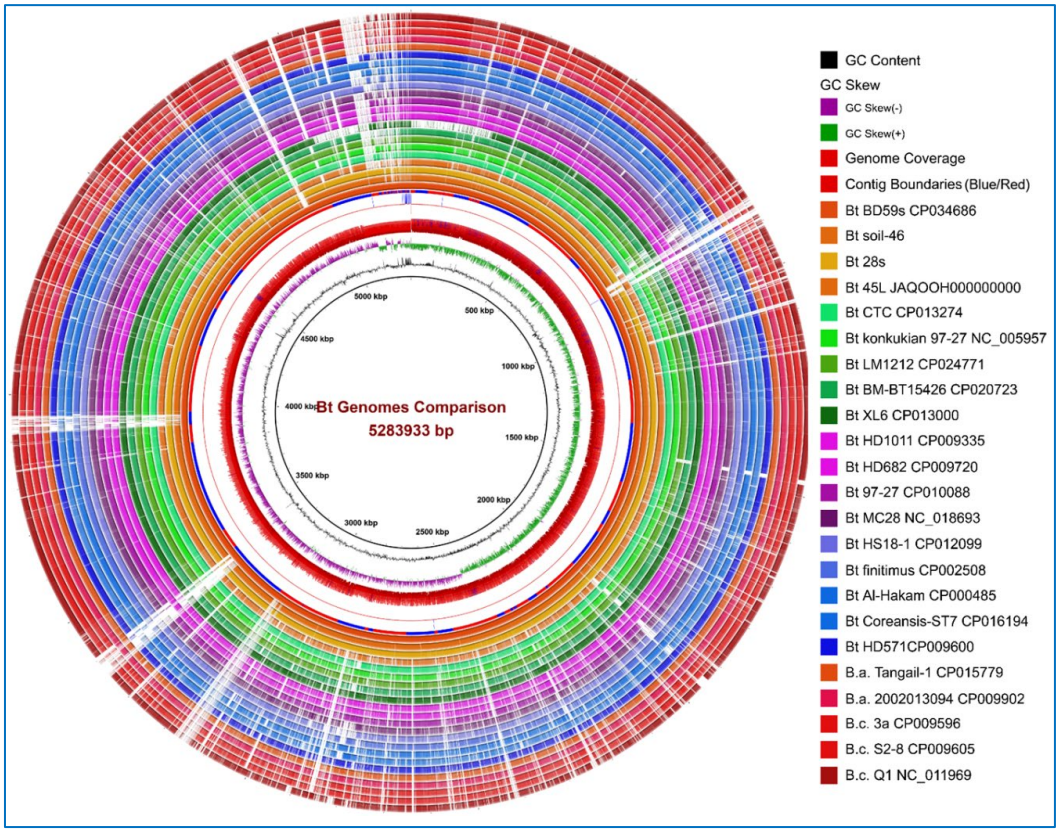



**Figure 6.3.4:** BRIG visualization shows BLAST comparisons against 18 genomes of *B. thuringiensis,* 2 *B. anthracis,* and 3 *B. cereus.* The innermost circles represent the GC content (black), GC skew (purple/green), genome coverage, and contig boundaries (alternating blue/red). After contig boundaries, circle 1 represents Bt BD59S, circles 2–18: *B. thuringiensis* (Bt)*,* circles 19–20: *B. anthracis* (B.a.), and circles 21–23: *B. cereus* (B.c.) genomes.

### 6.3.3.2 Multiple genome alignment by progressive mauve

Comparative genomics of 17 genome sequences of *B. thuringiensis* strains, 3 *B. cereus strains,* and 2 *B. anthracis* strains deposited in the NCBI GenBank was addressed using progressive mauve genome alignment.

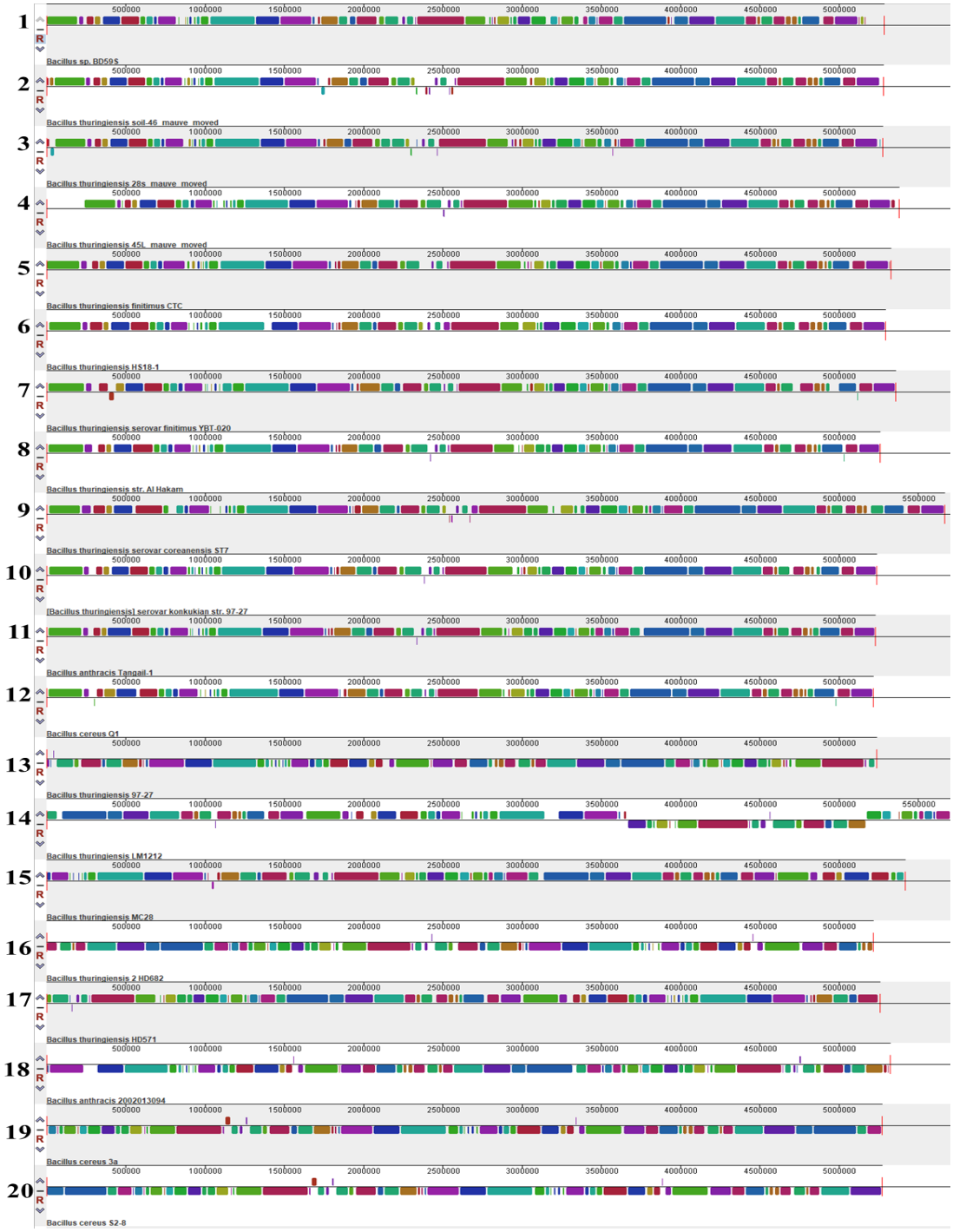



**Figure 6.3.5:** Comparative genomic map and comparisons of genome sequences of 17 *B. thuringiensis*, 3 *B. cereus,* and 2 *B. anthracis* strains. This alignment led to the formation of several extensive synteny-locally collinear blocks (LCBs). Each instance of the homologous section of the genome is shown as a distinct colored block, whereas dissimilar parts are represented as empty blocks or lines. In this case, the serial progression from higher to lower places, namely positions 1–12, exhibited consistent patterns, whereas positions 13–20 had distinct genetic patterns.

The alignment results showed a much larger synteny of colored locally collinear blocks (LCBs). Each syntenic placement of the homologous region of the genome is

represented as a unique colored block, while divergent regions are seen as an empty block or line (Figure 6.3.5). Based on this alignment, it was evident that the genomes of 10 Bt strains (BD59S, Soil-46, 28S, 45L, CTC (CP013274), HS18-1 (CP012099), *finitimus* Y020 (CP002508), Al-Hakam (CP000485), Coreansis-ST7 (CP016194), *konkukian* 97-27 (NC_005957), 1 Bc Q1 (NC_011969), and 1 Ba Tangain-1 (CP015779)) exhibited a common genomic organization pattern and homologous region compared to others (Figure 6.3.5).

#### 6.3.3.3 Phylogenomic study

To determine the identity and genotyping of indigenous Bt strains, NCBI microbial genome BLAST showed that all 4 indigenous Bt genomes are 97-98% similar to the whole genome sequences of thirteen *B. thuringiensis*, two *B. anthracis,* and three *B. cereus* genomes. To determine the phylogenetic relatedness, these strains were subjected to analysis by progressive mauve genome alignment (Darling *et al.* 2010), followed by Fig. v1.4.0 to view the phylogenetic relatedness amongst them. From phylogenomic tree analysis (Figure 6.3.6), it is seen that the genomes of three parasporal proteins synthesizing three different indigenous Bt strains, BD59S, Soil-46, and 28S, have a close relationship with the genome of Bt strain CTC (CP013274), and these four different genomes produce a separate node in the phylogenetic tree. Besides, among these 4 different Bt strains, three indigenous strains, Soil-46 and 28S again produce a separate clade, and from the taxa of the tree, it is clear that BD59S is slightly different from Soil-46 and 28S but they are from the same taxa. Besides, the single parasporal protein synthesizing another indigenous Bt genome 45L makes a separate clade with the three other genomes of Bt strain *konkukian* 97-27 (NC_005957), Bt 97-27 (CP010088), and Bt BMBT15426 (CP020723) (Figure 6.3.6). This study shows that the native Bt strains BD59S, Soil-46, 28S, and 45L are related to *B. thuringiensis* instead of *B. anthracis* and *B. cereus* (Figure 6.3.6). Besides, from this phylogenetic tree, it is very clear that the indigenous Bt strains BD59S, Soil-46, and 28S belong to *B. thuringiensis* serovar *finitimus*, and 45L belongs to the serovar *konkukian*.

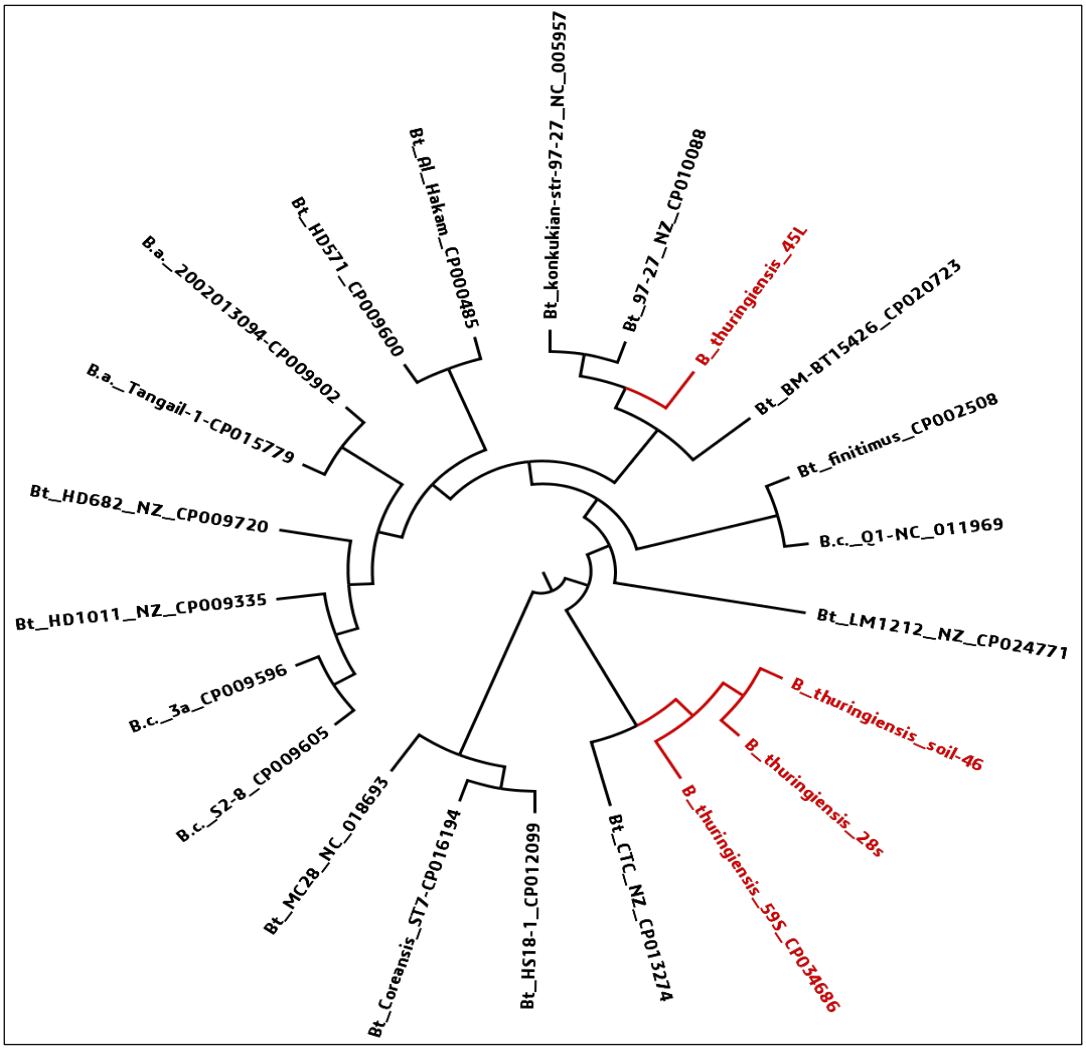



**Figure 6.3.6:** Phylogenetic relatedness of indigenous Bt genomes with the most similar complete genome sequence of *B. thuringiensis* (Bt), *B. anthracis* (Ba), and *B. cereus* (Bc) strains.

### 6.3.4 Identification and characterization of parasporal anticancer proteins from indigenous Bt genomes

We found parasporal proteins from native *B. thuringiensis* strains BD59S, 28S, and Soil-46 on SDS-PAGE. These proteins were 101, 86, and 76 kDa in size. Besides, the 103-kDa protein was found in 45L. To identify them, the NGS paired-end short-read fastq sequence files of each genome was aligned with all of the listed genbank accession number in the cry (http://www.lifesci.susx.ac.uk/home/Neil_Crickmore/Bt/; new website: https://www.bpprc-db.org/home/) and parasporin (https://www.fitc.pref.fukuoka.jp/parasporin/list.html) database. From the tablet visualization program, no aligned parasporin or cry gene was found to make any alignment with the NGS short reads of the indigenous Bt genome from BWA and the

Bowtie-2 alignment program. It indicated there are no known *cry* or *ps* genes in the indigenous Bt genomes.

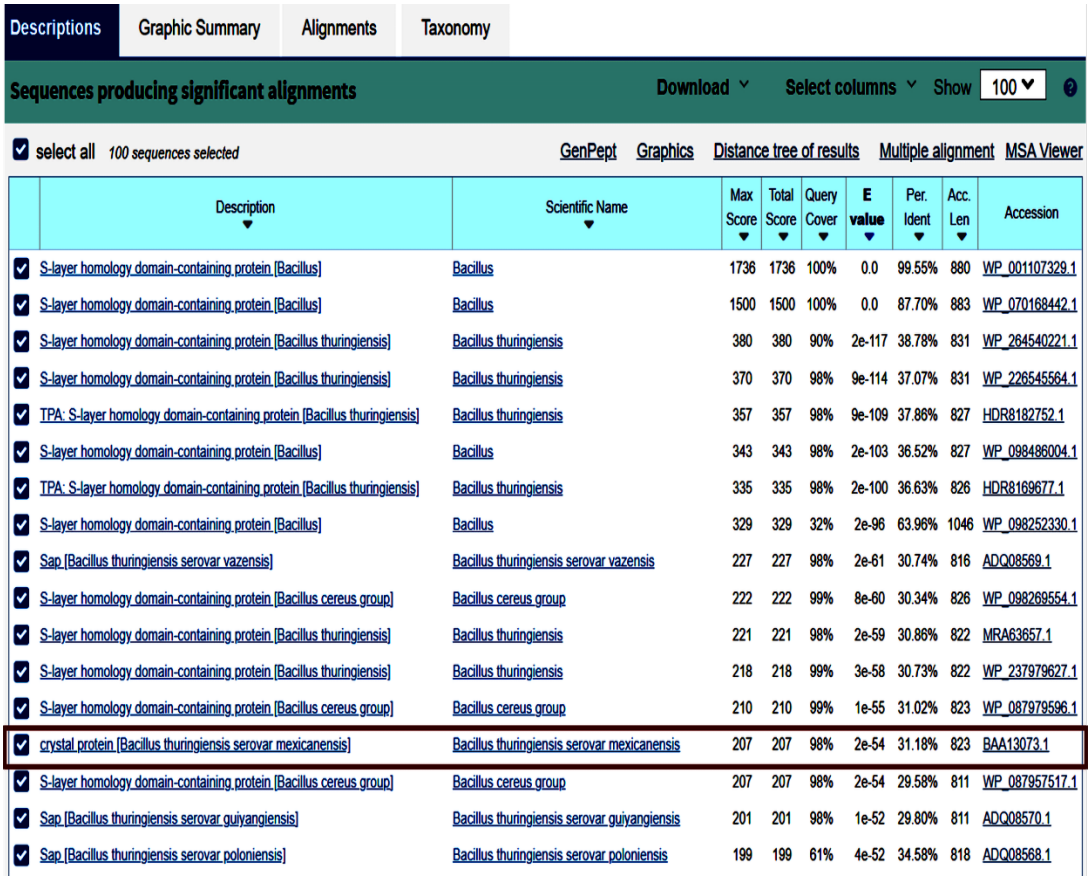
Then, from NCBI-PGAP and RAST annotated data, protein molecular weights ranging from 60 to 130 kDa were checked thoroughly using NCBI protein BLAST. From PGAP-annotated data instead of Cry, a few protein sequences with the desired size were identified. Among them, two S-layer proteins (SLPs) with different molecular weights were found. From the Uniport BLAST and InterPro protein signature databases scan, it was found that each of BD59S, 28S, and Soil-46 had two different types of S-layer homology-domain (SLH)-containing protein (SLP); of them, one was an 814 amino acid long 86 kDa surface array protein (Sap), and another was an 859 amino acid long 91 kDa extractable antigen-1 (EA1). Besides, a 73.54 kDa hypothetical protein was also found in each of these genomes. Besides, from the 45L genome sequence, two different molecular weight sizes of S-layer proteins (SLPs) were found. The SLPs were an 883 amino acid long (93.29 kDa) SLP and a 900 amino acid long (103.11 kDa) SLP. From the InterPro scan, it was found that the 103 kDa SLP had low-level similarities with EA1 and the 93 kDa SLP also had low-level similarities with Sap. Each of these protein molecules was expected to produce parasporal inclusions of the respective genome. Thus, further bioinformatic analysis was performed to confirm their inclusion forming capabilities.

#### 6.3.4.1 NCBI BLAST identification of indigenous S-layer proteins

From NCBI protein BLAST, it was found that the 86 kDa S-layer protein, Sap, of *B. thuringiensis* strains BD59S, 28S, and Soil-46 had 100.00% to 78% sequence identity with the Sap of another *B. thuringiensis, B. cereus, B. paranthracis, B. anthracis, B. mycoids,* and *B. cytotoxicus.* On the other hand, from NCBI protein BLAST, it was found that another S-layer protein, EA1 of *B. thuringiensis* strains BD59S, 28S, and Soil-46, had 99.88% to 88.52% sequence identity with the EA1 of another *B. thuringiensis, B. cereus, B. paranthracis, B. anthracis, B. mycoids, B. pacificus,* and *B. wiedmannii.* Besides, from the NCBI protein BLAST result of 73.54 kDa protein of 3 *B. thuringiensis* strains BD59S, Soil-46, and 28S, it was found that the protein sequence

had a 100.00% to 96% sequence identity with the hypothetical proteins of others, Bt and *Bacillus sp*. bacterium.

From NCBI protein BLAST, it was found that the 883 amino acid-long 93.29 kDa S-layer protein of *B. thuringiensis* strains 45L had 99.55% to 30% sequence identity with the S-layer homology domain-containing protein of *B. thuringiensis* and *B. cereus* group bacteria (Figure 6.3.7). Besides, 93.29 kDa S-layer protein also had 31% sequence identity with the crystal protein of Bt mexicanensis (BAA13073.1), and from the protein feature study, we found that it was a parasporal inclusion forming S-layer protein. On the other hand, from NCBI protein BLAST, it was found that the 900 amino acid long 103.11 kDa S-layer protein of *B. thuringiensis* strains 45L had 99.56% to 92.48% sequence similarity with the S-layer homology domain-containing protein of *Bacillus sp*. and *B. cereus* bacteria (Figure 6.3.8). From the NCBI protein BLAST, an interesting finding was observed: the 99.56% identical protein sequence was matched with a parasporal protein of *Streptococcus pneumoniae* (COE74891.1) (Figure 6.3.8). It was submitted by the Welcome Trust Sanger Institute of the United Kingdom but no additional information was obtained about this parasporal protein of *S. pneumoniae.*



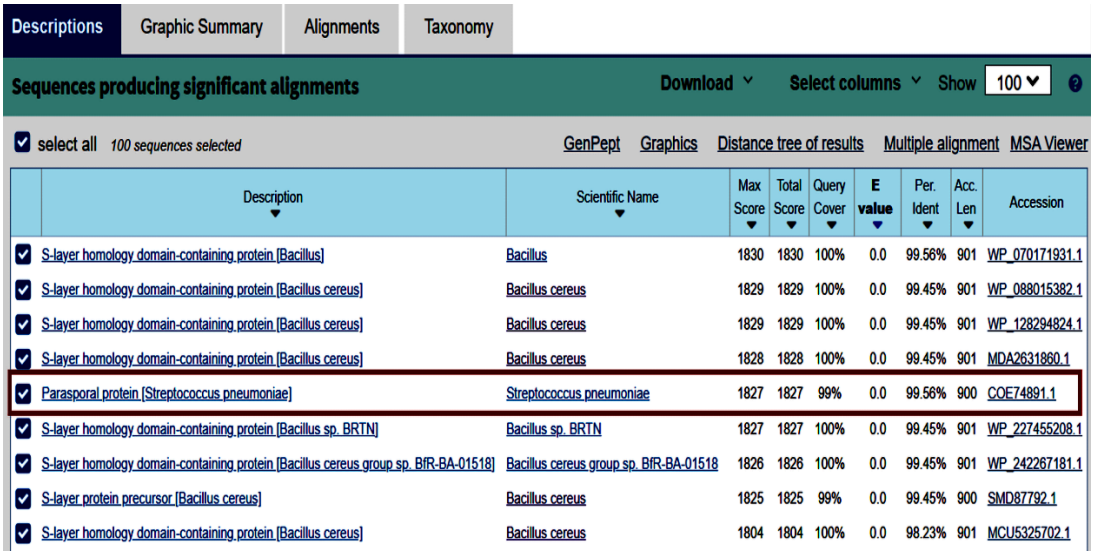
Descriptions | Graphic Summary | Alignments | Taxonomy

Sequences producing significant alignments — Download | Select columns | Show 100

select all 100 sequences selected — GenPept | Graphics | Distance tree of results | Multiple alignment | MSA Viewer

| Description | Scientific Name | Max Score | Total Score | Query Cover | E value | Per. Ident | Acc. Len | Accession |
|---|---|---|---|---|---|---|---|---|
| S-layer homology domain-containing protein [Bacillus] | Bacillus | 1736 | 1736 | 100% | 0.0 | 99.55% | 880 | WP_001107329.1 |
| S-layer homology domain-containing protein [Bacillus] | Bacillus | 1500 | 1500 | 100% | 0.0 | 87.70% | 883 | WP_070168442.1 |
| S-layer homology domain-containing protein [Bacillus thuringiensis] | Bacillus thuringiensis | 380 | 380 | 90% | 2e-117 | 38.78% | 831 | WP_264540221.1 |
| S-layer homology domain-containing protein [Bacillus thuringiensis] | Bacillus thuringiensis | 370 | 370 | 98% | 9e-114 | 37.07% | 831 | WP_226545564.1 |
| TPA: S-layer homology domain-containing protein [Bacillus thuringiensis] | Bacillus thuringiensis | 357 | 357 | 98% | 9e-109 | 37.86% | 827 | HDR8182752.1 |
| S-layer homology domain-containing protein [Bacillus] | Bacillus | 343 | 343 | 98% | 2e-103 | 36.52% | 827 | WP_098486004.1 |
| TPA: S-layer homology domain-containing protein [Bacillus thuringiensis] | Bacillus thuringiensis | 335 | 335 | 98% | 2e-100 | 36.63% | 826 | HDR8169677.1 |
| S-layer homology domain-containing protein [Bacillus] | Bacillus | 329 | 329 | 32% | 2e-96 | 63.96% | 1046 | WP_098252330.1 |
| Sap [Bacillus thuringiensis serovar vazensis] | Bacillus thuringiensis serovar vazensis | 227 | 227 | 98% | 2e-61 | 30.74% | 816 | ADQ08569.1 |
| S-layer homology domain-containing protein [Bacillus cereus group] | Bacillus cereus group | 222 | 222 | 99% | 8e-60 | 30.34% | 826 | WP_098269554.1 |
| S-layer homology domain-containing protein [Bacillus thuringiensis] | Bacillus thuringiensis | 221 | 221 | 98% | 2e-59 | 30.86% | 822 | MRA63657.1 |
| S-layer homology domain-containing protein [Bacillus thuringiensis] | Bacillus thuringiensis | 218 | 218 | 99% | 3e-58 | 30.73% | 822 | WP_237979627.1 |
| S-layer homology domain-containing protein [Bacillus cereus group] | Bacillus cereus group | 210 | 210 | 99% | 1e-55 | 31.02% | 823 | WP_087979596.1 |
| crystal protein [Bacillus thuringiensis serovar mexicanensis] | Bacillus thuringiensis serovar mexicanensis | 207 | 207 | 98% | 2e-54 | 31.18% | 823 | BAA13073.1 |
| S-layer homology domain-containing protein [Bacillus cereus group] | Bacillus cereus group | 207 | 207 | 98% | 2e-54 | 29.58% | 811 | WP_087957517.1 |
| Sap [Bacillus thuringiensis serovar guiyangiensis] | Bacillus thuringiensis serovar guiyangiensis | 201 | 201 | 98% | 1e-52 | 29.80% | 811 | ADQ08570.1 |
| Sap [Bacillus thuringiensis serovar poloniensis] | Bacillus thuringiensis serovar poloniensis | 199 | 199 | 61% | 4e-52 | 34.58% | 818 | ADQ08568.1 |

**Figure 6.3.7:** NCBI protein BLAST result of 883 amino acid-long 93.29 kDa S-layer protein of *B. thuringiensis* strains 45L.



Descriptions | Graphic Summary | Alignments | Taxonomy

Sequences producing significant alignments — Download | Select columns | Show 100

select all 100 sequences selected — GenPept | Graphics | Distance tree of results | Multiple alignment | MSA Viewer

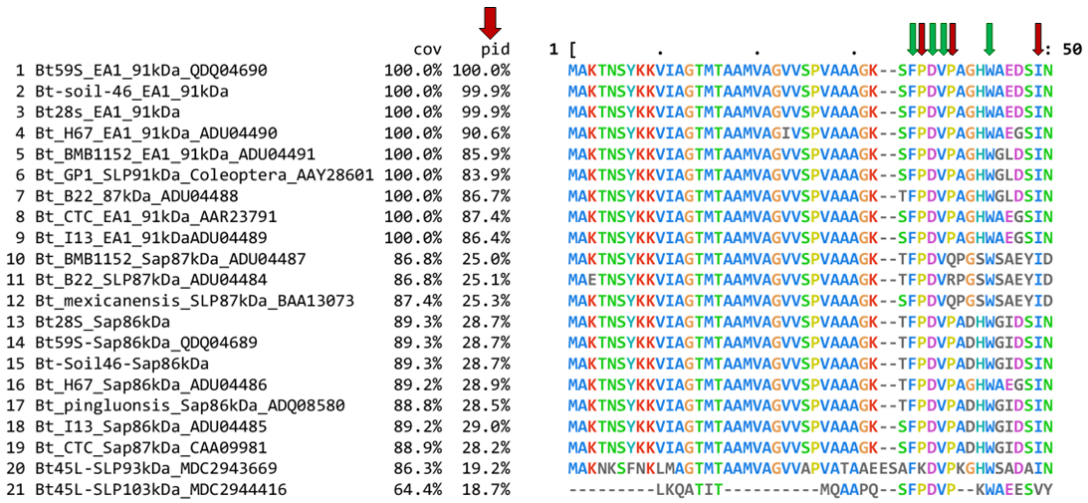
| Description | Scientific Name | Max Score | Total Score | Query Cover | E value | Per. Ident | Acc. Len | Accession |
|---|---|---|---|---|---|---|---|---|
| S-layer homology domain-containing protein [Bacillus] | Bacillus | 1830 | 1830 | 100% | 0.0 | 99.56% | 901 | WP_070171931.1 |
| S-layer homology domain-containing protein [Bacillus cereus] | Bacillus cereus | 1829 | 1829 | 100% | 0.0 | 99.45% | 901 | WP_088015382.1 |
| S-layer homology domain-containing protein [Bacillus cereus] | Bacillus cereus | 1829 | 1829 | 100% | 0.0 | 99.45% | 901 | WP_128294824.1 |
| S-layer homology domain-containing protein [Bacillus cereus] | Bacillus cereus | 1828 | 1828 | 100% | 0.0 | 99.45% | 901 | MDA2631860.1 |
| Parasporal protein [Streptococcus pneumoniae] | Streptococcus pneumoniae | 1827 | 1827 | 99% | 0.0 | 99.56% | 900 | COE74891.1 |
| S-layer homology domain-containing protein [Bacillus sp. BRTN] | Bacillus sp. BRTN | 1827 | 1827 | 100% | 0.0 | 99.45% | 901 | WP_227455208.1 |
| S-layer homology domain-containing protein [Bacillus cereus group sp. BfR-BA-01518] | Bacillus cereus group sp. BfR-BA-01518 | 1826 | 1826 | 100% | 0.0 | 99.45% | 901 | WP_242267181.1 |
| S-layer protein precursor [Bacillus cereus] | Bacillus cereus | 1825 | 1825 | 99% | 0.0 | 99.45% | 900 | SMD87792.1 |
| S-layer homology domain-containing protein [Bacillus cereus] | Bacillus cereus | 1804 | 1804 | 100% | 0.0 | 98.23% | 901 | MCU5325702.1 |

**Figure 6.3.8:** NCBI protein BLAST result of 900 amino acid-long 103.11 kDa S-layer protein of *B. thuringiensis* strains 45L.

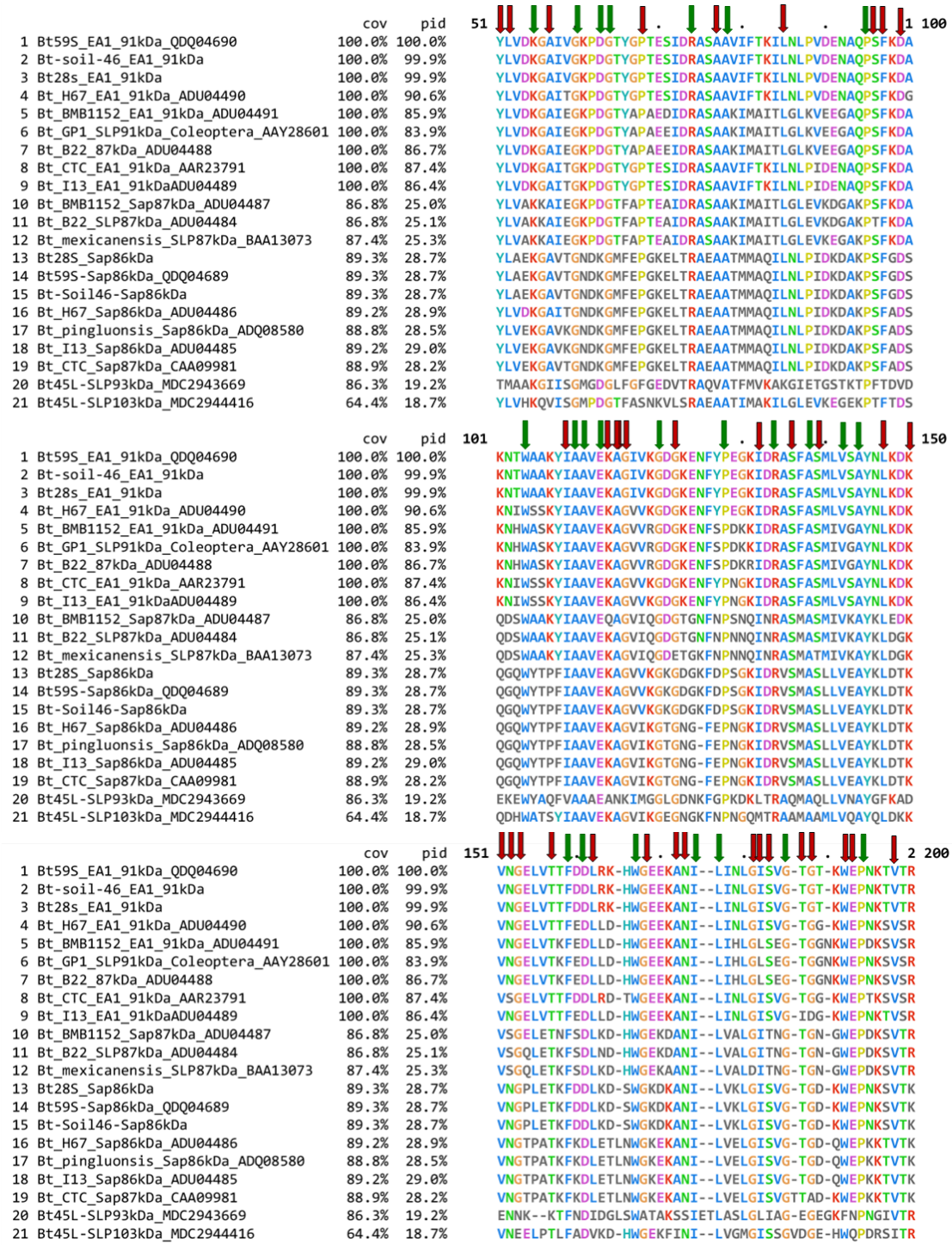
### 6.3.4.2 Multiple sequence alignments between indigenous and parasporal inclusion forming S-layer proteins

From various published literature, it was found that both the S-layer protein Sap and EA1 were able to produce round or oval-shaped parasporal inclusions in many *B. thuringiensis* strains. Hence, all of the previously identified and experimentally proved parasporal inclusion forming S-layer protein sequence data (EA1 and Sap) were collected from the NCBI and respective publications. Then, all of these S-layer protein sequences were compared with the SLP of indigenous *B. thuringiensis* genomes by MAFFT multiple sequence alignment.

The MSA of indigenous and parasporal inclusion forming S-layer proteins (Figure 6.3.9) demonstrates the sequence similarities between our S-layer protein and 13 other parasporal S-layer proteins. In the alignment, the initial 30 sequences are signal peptide regions and amino acid sequences of signal peptides are homogeneous. The conserved SLH homology domain is located between 33 and 201, and the rest are crystallization or assembly domains. The conserved SLH domain of the parasporal Sap, EA1, and other types of parasporal S-layer proteins differed significantly from the known classes, exhibiting a low degree of similarity (<30%). The green arrow indicates highly conserved amino acid sequences present in all SLPs; the red arrow indicates conserved amino acid sequences present in most of the SLPs except 1 to 5. From MSA, it was found that SLPs are genealogically heterogeneous.




Continued


cov pid 51 1 100
1 Bt59S_EA1_91kDa_QDQ04690 100.0% 100.0% YLVDKGAIVGKPDGTYGPTESIDRASAAVIFTKILNLPVDENAQPSFKDA
2 Bt-soil-46_EA1_91kDa 100.0% 99.9% YLVDKGAIVGKPDGTYGPTESIDRASAAVIFTKILNLPVDENAQPSFKDA
3 Bt28s_EA1_91kDa 100.0% 99.9% YLVDKGAIVGKPDGTYGPTESIDRASAAVIFTKILNLPVDENAQPSFKDA
4 Bt_H67_EA1_91kDa_ADU04490 100.0% 90.6% YLVDKGAITGKPDGTYGPTESIDRASAAVIFTKILNLPVDENAQPSFKDG
5 Bt_BMB1152_EA1_91kDa_ADU04491 100.0% 85.9% YLVDKGAIEGKPDGTYAPAEDIDRASAAKIMAITLGLKVEEGAQPSFKDA
6 Bt_GP1_SLP91kDa_Coleoptera_AAY28601 100.0% 83.9% YLVDKGAIEGKPDGTYAPAEEIDRASAAKIMAITLGLKVEEGAQPSFKDA
7 Bt_B22_87kDa_ADU04488 100.0% 86.7% YLVDKGAIEGKPDGTYAPAEEIDRASAAKIMAITLGLKVEEGAQPSFKDA
8 Bt_CTC_EA1_91kDa_AAR23791 100.0% 87.4% YLVDKGAITGKPDGTYGPTESIDRASAAVIFTKILNLPIDENAQPSFKDA
9 Bt_I13_EA1_91kDaADU04489 100.0% 86.4% YLVDKGAITGKPDGTYGPTESIDRASAAVIFTKILNLPIDENAQPSFKDA
10 Bt_BMB1152_Sap87kDa_ADU04487 86.8% 25.0% YLVAKKAIEGKPDGTFAPTEAIDRASAAKIMAITLGLEVKDGAKPSFKDA
11 Bt_B22_SLP87kDa_ADU04484 86.8% 25.1% YLVAKKAIEGKPDGTFAPTEAIDRASAAKIMAITLGLEVKDGAKPTFKDA
12 Bt_mexicanensis_SLP87kDa_BAA13073 87.4% 25.3% YLVAKKAIEGKPDGTFAPTEAIDRASAAKIMAITLGLEVKEGAKPSFKDA
13 Bt28S_Sap86kDa 89.3% 28.7% YLAEKGAVTGNDKGMFEPGKELTRAEAATMMAQILNLPIDKDAKPSFGDS
14 Bt59S-Sap86kDa_QDQ04689 89.3% 28.7% YLAEKGAVTGNDKGMFEPGKELTRAEAATMMAQILNLPIDKDAKPSFGDS
15 Bt-Soil46-Sap86kDa 89.3% 28.7% YLAEKGAVTGNDKGMFEPGKELTRAEAATMMAQILNLPIDKDAKPSFGDS
16 Bt_H67_Sap86kDa_ADU04486 89.2% 28.9% YLVDKGAITGNDKGMFEPGKELTRAEAATMMAQILNLPIDKDAKPSFADS
17 Bt_pingluonsis_Sap86kDa_ADQ08580 88.8% 28.5% YLVEKGAVKGNDKGMFEPGKELTRAEAATMMAQILNLPIDKDAKPSFADS
18 Bt_I13_Sap86kDa_ADU04485 89.2% 29.0% YLVEKGAVKGNDKGMFEPGKELTRAEAATMMAQILNLPIDKDAKPSFADS
19 Bt_CTC_Sap87kDa_CAA09981 88.9% 28.2% YLVEKGAVTGNDKGMFEPGKELTRAEAATMMAQILNLPIDKDAKPSFADS
20 Bt45L-SLP93kDa_MDC2943669 86.3% 19.2% TMAAKGIISGMGDGLFGFGEDVTRAQVATFMVKAKGIETGSTKTPFTDVD
21 Bt45L-SLP103kDa_MDC2944416 64.4% 18.7% YLVHKQVISGMPDGTFASNKVLSRAEAATIMAKILGLEVKEGEKPTFTDS
cov pid 101 150
1 Bt59S_EA1_91kDa_QDQ04690 100.0% 100.0% KNTWAAKYIAAVEKAGIVKGDGKENFYPEGKIDRASFASMLVSAYNLKDK
2 Bt-soil-46_EA1_91kDa 100.0% 99.9% KNTWAAKYIAAVEKAGIVKGDGKENFYPEGKIDRASFASMLVSAYNLKDK
3 Bt28s_EA1_91kDa 100.0% 99.9% KNTWAAKYIAAVEKAGIVKGDGKENFYPEGKIDRASFASMLVSAYNLKDK
4 Bt_H67_EA1_91kDa_ADU04490 100.0% 90.6% KNIWSSKYIAAVEKAGVVKGDGKENFYPEGKIDRASFASMLVSAYNLKDK
5 Bt_BMB1152_EA1_91kDa_ADU04491 100.0% 85.9% KNHWASKYIAAVEKAGVVRGDGKENFSPDKKIDRASFASMIVGAYNLKDK
6 Bt_GP1_SLP91kDa_Coleoptera_AAY28601 100.0% 83.9% KNHWASKYIAAVEKAGVVRGDGKENFSPDKKIDRASFASMIVGAYNLKDK
7 Bt_B22_87kDa_ADU04488 100.0% 86.7% KNHWASKYIAAVEKAGVVRGDGKENFSPDKRIDRASFASMIVGAYNLKDK
8 Bt_CTC_EA1_91kDa_AAR23791 100.0% 87.4% KNIWSSKYIAAVEKAGVVKGDGKENFYPNGKIDRASFASMLVSAYNLKDK
9 Bt_I13_EA1_91kDaADU04489 100.0% 86.4% KNIWSSKYIAAVEKAGVVKGDGKENFYPNGKIDRASFASMLVSAYNLKDK
10 Bt_BMB1152_Sap87kDa_ADU04487 86.8% 25.0% QDSWAAKYIAAVEQAGVIQGDGTGNFNPSNQINRASMASMIVKAYKLEDK
11 Bt_B22_SLP87kDa_ADU04484 86.8% 25.1% QDSWAAKYIAAVEKAGVIQGDGTGNFNPNNQINRASMASMIVKAYKLDGK
12 Bt_mexicanensis_SLP87kDa_BAA13073 87.4% 25.3% QDSWAAKYIAAVEKAGVIQGDETGKFNPNNQINRASMATMIVKAYKLDGK
13 Bt28S_Sap86kDa 89.3% 28.7% QGQWYTPFIAAVEKAGVVKGKGDGKFDPSGKIDRVSMASLLVEAYKLDTK
14 Bt59S-Sap86kDa_QDQ04689 89.3% 28.7% QGQWYTPFIAAVEKAGVVKGKGDGKFDPSGKIDRVSMASLLVEAYKLDTK
15 Bt-Soil46-Sap86kDa 89.3% 28.7% QGQWYTPFIAAVEKAGVVKGKGDGKFDPSGKIDRVSMASLLVEAYKLDTK
16 Bt_H67_Sap86kDa_ADU04486 89.2% 28.9% QGQWYTPFIAAVEKAGVIKGTGNG-FEPNGKIDRVSMASLLVEAYKLDTK
17 Bt_pingluonsis_Sap86kDa_ADQ08580 88.8% 28.5% QGQWYTPFIAAVEKAGVIKGTGNG-FEPNGKIDRVSMASLLVEAYKLDTK
18 Bt_I13_Sap86kDa_ADU04485 89.2% 29.0% QGQWYTPFIAAVEKAGVIKGTGNG-FEPNGKIDRVSMASLLVEAYKLDTK
19 Bt_CTC_Sap87kDa_CAA09981 88.9% 28.2% QGQWYTPFIAAVEKAGVIKGTGNG-FEPNGKIDRVSMASLLVEAYKLDTK
20 Bt45L-SLP93kDa_MDC2943669 86.3% 19.2% EKEWYAQFVAAAEANKIMGGLGDNKFGPKDKLTRAQMAQLLVNAYGFKAD
21 Bt45L-SLP103kDa_MDC2944416 64.4% 18.7% QDHWATSYIAAVEKAGVIKGEGNGKFNPNGQMTRAAMAAMLVQAYQLDKK
cov pid 151 2 200
1 Bt59S_EA1_91kDa_QDQ04690 100.0% 100.0% VNGELVTTFDDLRK-HWGEEKANI--LINLGISVG-TGT-KWEPNKTVTR
2 Bt-soil-46_EA1_91kDa 100.0% 99.9% VNGELVTTFDDLRK-HWGEEKANI--LINLGISVG-TGT-KWEPNKTVTR
3 Bt28s_EA1_91kDa 100.0% 99.9% VNGELVTTFDDLRK-HWGEEKANI--LINLGISVG-TGT-KWEPNKTVTR
4 Bt_H67_EA1_91kDa_ADU04490 100.0% 90.6% VNGELVTTFEDLLD-HWGEEKANI--LINLGISVG-TGG-KWEPNKSVSR
5 Bt_BMB1152_EA1_91kDa_ADU04491 100.0% 85.9% VNGELVTKFEDLLD-HWGEEKANI--LIHLGLSEG-TGGNKWEPDKSVSR
6 Bt_GP1_SLP91kDa_Coleoptera_AAY28601 100.0% 83.9% VNGELVTKFEDLLD-HWGEEKANI--LIHLGLSEG-TGGNKWEPNKSVSR
7 Bt_B22_87kDa_ADU04488 100.0% 86.7% VNGELVTKFEDLLD-HWGEEKANI--LIHLGLSEG-TGGNKWEPNKSVSR
8 Bt_CTC_EA1_91kDa_AAR23791 100.0% 87.4% VSGELVTTFDDLRD-TWGEEKANI--LINLGISVG-TGG-KWEPTKSVSR
9 Bt_I13_EA1_91kDaADU04489 100.0% 86.4% VNGELVTTFEDLLD-HWGEEKANI--LINLGISVG-IDG-KWEPNKTVSR
10 Bt_BMB1152_Sap87kDa_ADU04487 86.8% 25.0% VSGELETNFSDLKD-HWGEKDANI--LVALGITNG-TGN-GWEPDKSVTR
11 Bt_B22_SLP87kDa_ADU04484 86.8% 25.1% VSGQLETKFSDLND-HWGEKDANI--LVALGITNG-TGN-GWEPDKSVTR
12 Bt_mexicanensis_SLP87kDa_BAA13073 87.4% 25.3% VSGQLETKFSDLKD-HWGEKAANI--LVALDITNG-TGN-GWEPDKSVTR
13 Bt28S_Sap86kDa 89.3% 28.7% VNGPLETKFDDLKD-SWGKDKANI--LVKLGISVG-TGD-KWEPNKSVTK
14 Bt59S-Sap86kDa_QDQ04689 89.3% 28.7% VNGPLETKFDDLKD-SWGKDKANI--LVKLGISVG-TGD-KWEPNKSVTK
15 Bt-Soil46-Sap86kDa 89.3% 28.7% VNGPLETKFDDLKD-SWGKDKANI--LVKLGISVG-TGD-KWEPNKSVTK
16 Bt_H67_Sap86kDa_ADU04486 89.2% 28.9% VNGTPATKFKDLETLNWGKEKANI--LVELGISVG-TGD-QWEPKKTVTK
17 Bt_pingluonsis_Sap86kDa_ADQ08580 88.8% 28.5% VNGTPATKFKDLETLNWGKEKANI--LVELGISVG-TGD-QWEPKKTVTK
18 Bt_I13_Sap86kDa_ADU04485 89.2% 29.0% VNGTPATKFKDLETLNWGKEKANI--LVELGISVG-TGD-QWEPKKTVTK
19 Bt_CTC_Sap87kDa_CAA09981 88.9% 28.2% VNGTPATKFKDLETLNWGKEKANI--LVELGISVGTTAD-KWEPKKTVTK
20 Bt45L-SLP93kDa_MDC2943669 86.3% 19.2% ENNK--KTFNDIDGLSWATAKSSIETLASLGLIAG-EGEGKFNPNGIVTR
21 Bt45L-SLP103kDa_MDC2944416 64.4% 18.7% VNEELPTLFADVKD-HWGEKFINI--LVGMGISSGVDGE-HWQPDRSITR

Continued

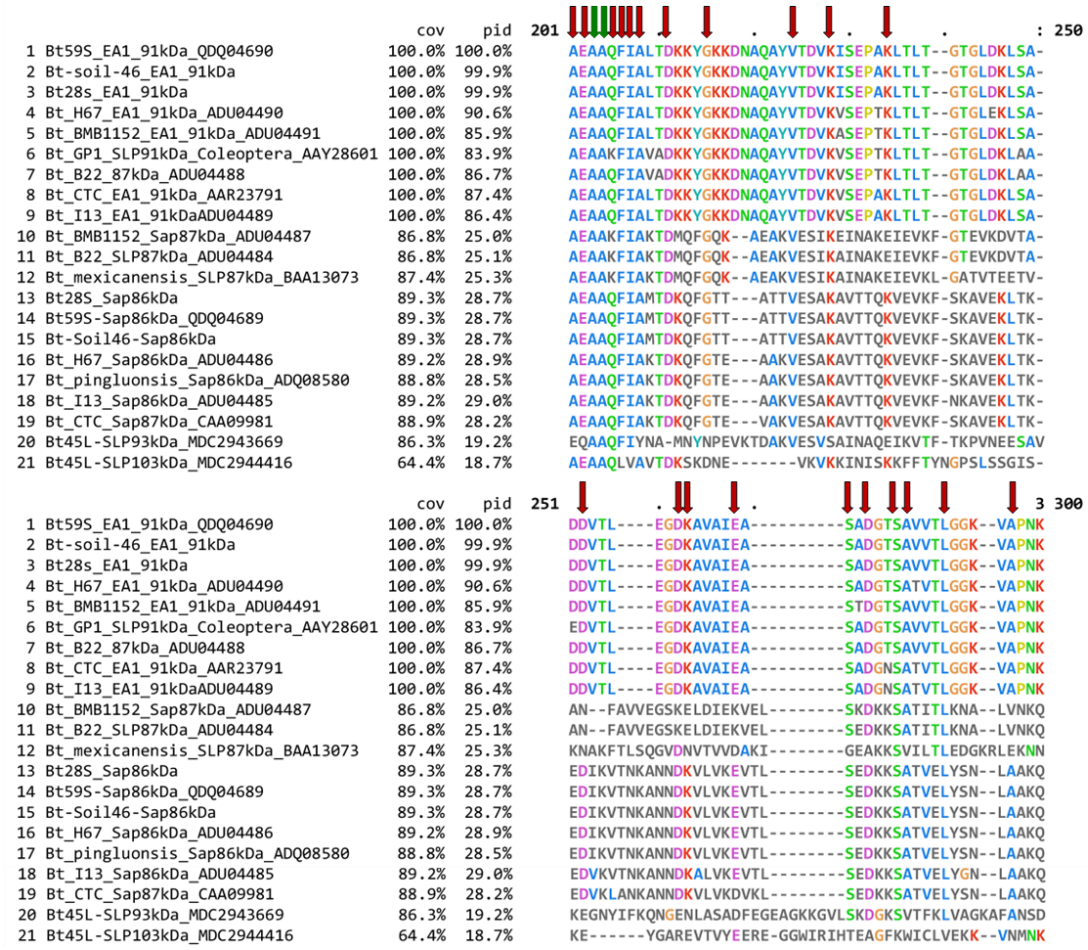
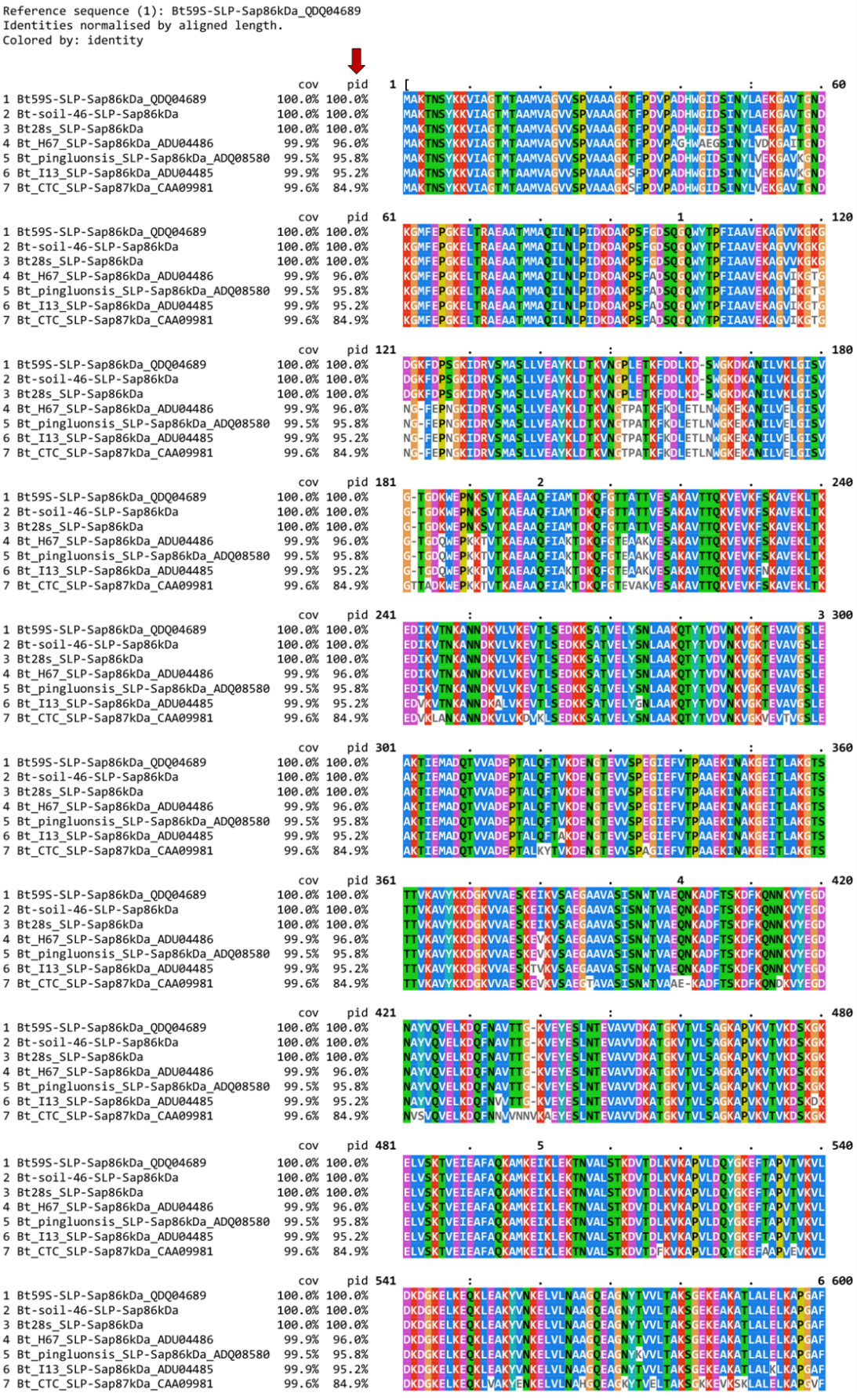



**Figure 6.3.9 :** MSA of indigenous and parasporal inclusion forming S-layer protein of Bt strains. (pid means percent identity. The green arrow indicates highly conserved amino acid sequences present in all SLPs; the red arrow indicates conserved amino acid sequences present in most of the SLPs except 1 to 5 amino acids.)

***MSA and percent identity of 86 kDa Sap of BD59S, Soil-46, and 28S with parasporal inclusion forming Sap of other Bt strains:***

From MAFFT alignments and the MView percent identity matrix, it was found that the 86 kDa S-layer protein Sap of 3 indigenous Bt strains had 96% to 85% amino acid sequence similarities with previously reported 86 kDa four parasporal S-layer protein Sap (ADU04486.1, ADQ08580.1, ADU04485.1, and CAA09981.1) of *B. thuringiensis* strains (Figure 6.3.10). Based on the MSA and the level of percent identity, it is reasonable to infer that there is a strong correlation between the 86 kDa Sap with the parasporal inclusion forming Sap of other Bt strains. In addition, based on similarities of the size of the 86 kDa Sap and the SDS-PAGE observed 86 kDa parasporal protein band, it can be said that the 86 kDa Sap of the indigenous Bt strains BD59S, 28S, and Soil-46 is the 86 kDa parasporal inclusion-forming protein of those strains.


Reference sequence (1): Bt59S-SLP-Sap86kDa_QDQ04689
Identities normalised by aligned length.
Colored by: identity
cov pid
1 Bt59S-SLP-Sap86kDa_QDQ04689 100.0% 100.0%
2 Bt-soil-46-SLP-Sap86kDa 100.0% 100.0%
3 Bt28s_SLP-Sap86kDa 100.0% 100.0%
4 Bt_H67_SLP-Sap86kDa_ADU04486 99.9% 96.0%
5 Bt_pingluonsis_SLP-Sap86kDa_ADQ08580 99.5% 95.8%
6 Bt_I13_SLP-Sap86kDa_ADU04485 99.9% 95.2%
7 Bt_CTC_SLP-Sap87kDa_CAA09981 99.6% 84.9%

Continued

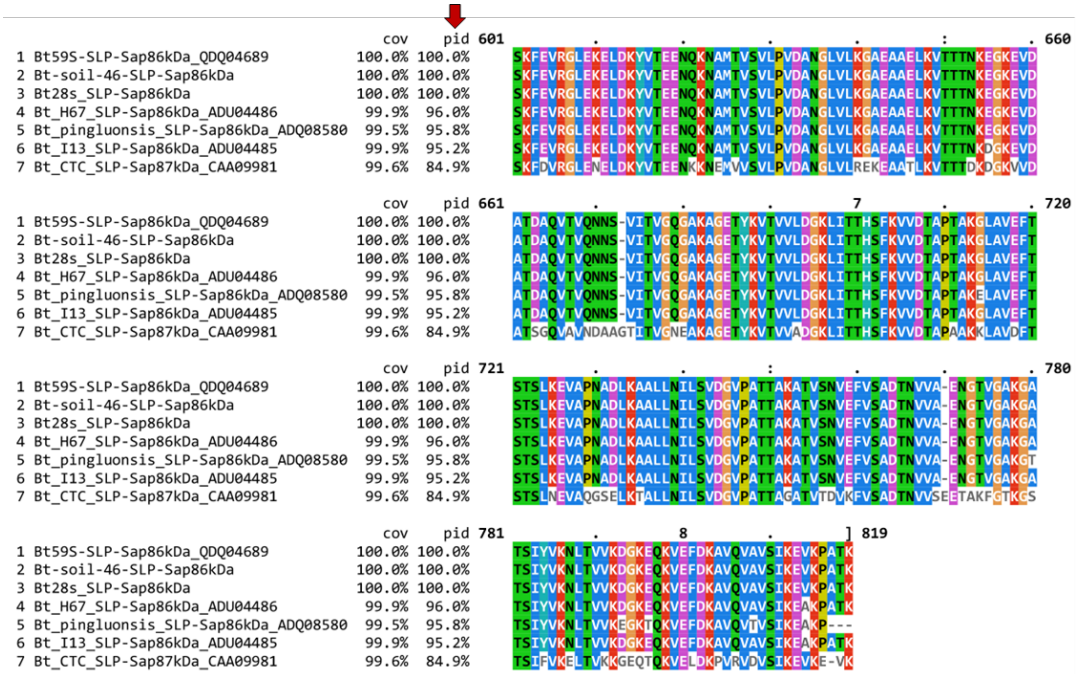
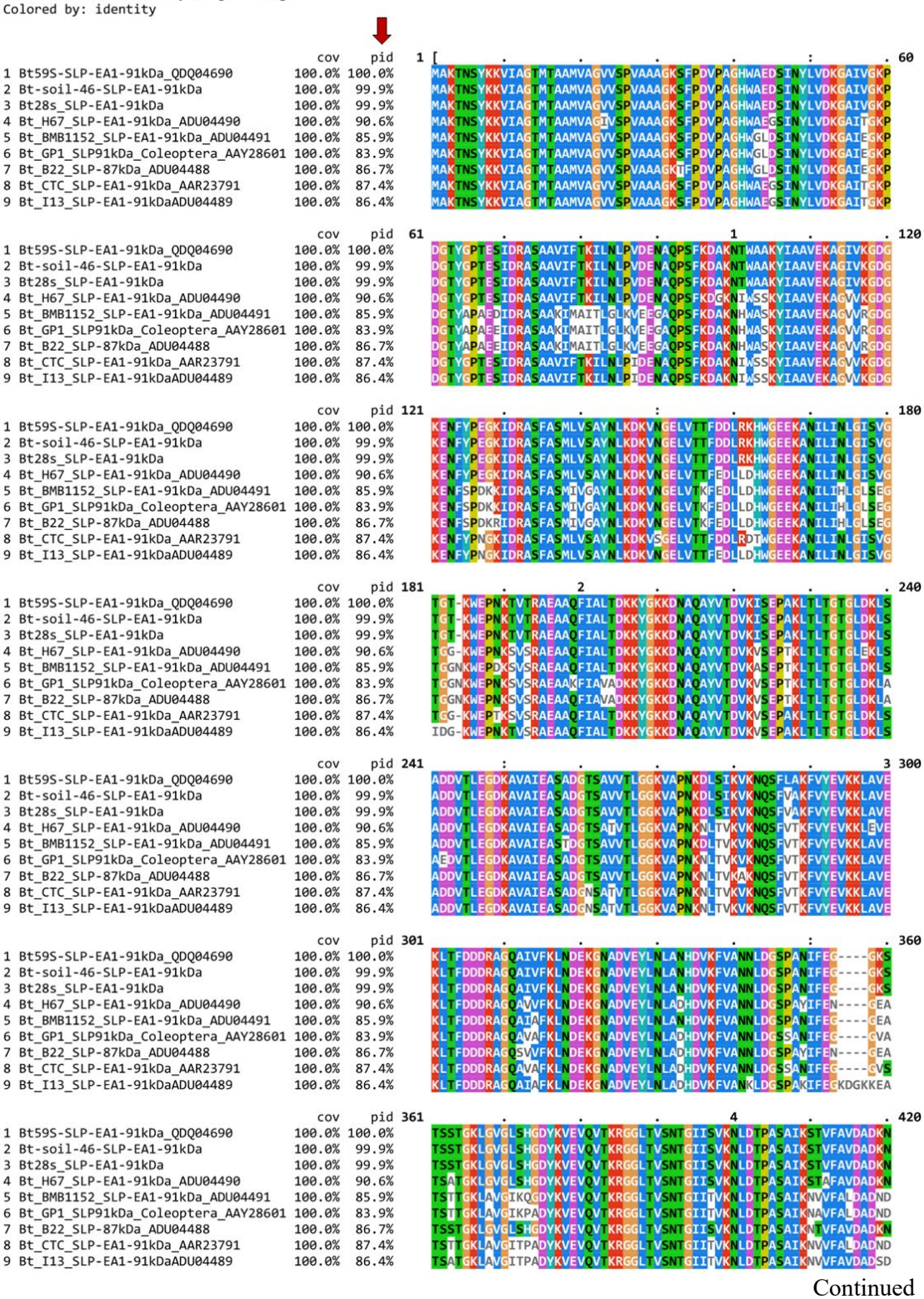



**Figure 6.3.10:** MSA and percent identity of 86 kDa Sap of *B. thuringiensis* strains BD59S, Soil-46, and 28S with parasporal inclusion forming Sap of other Bt strains. (pid means percent identity)

***MSA and percent identity of 91 kDa EA1 of BD59S, Soil-46, and 28S with parasporal inclusion forming EA1 of other Bt strains:***

The 91 kDa S-layer protein EA1 of three indigenous strains had 90% to 84% sequence similarities with previously reported 91 kDa 6 parasporal S-layer protein EA1 of different *B. thuringiensis* strains (Figure 6.3.11) (NCBI protein accession AAR23791.1, ADU04488.1, ADU04489.1, ADU04490.1, ADU04491.1, and AAY28601.1). Based on the MSA and the level of percent identity, it is reasonable to infer that there is a strong correlation between the 91 kDa EA-1 with the parasporal inclusion forming EA-1 of other Bt strains. In addition, based on similarities of the size of the 91 kDa EA-1 and the SDS-PAGE observed ~100 kDa parasporal protein band, it can be said that the 91 kDa EA-1 of the indigenous Bt strains BD59S, 28S, and Soil-46 could be the 91 kDa parasporal inclusion-forming protein of those strains.


Reference sequence (1): Bt59S-SLP-EA1-91kDa_QDQ04690
Identities normalised by aligned length.
Colored by: identity
cov pid
1 Bt59S-SLP-EA1-91kDa_QDQ04690 100.0% 100.0%
2 Bt-soil-46-SLP-EA1-91kDa 100.0% 99.9%
3 Bt28s_SLP-EA1-91kDa 100.0% 99.9%
4 Bt_H67_SLP-EA1-91kDa_ADU04490 100.0% 90.6%
5 Bt_BMB1152_SLP-EA1-91kDa_ADU04491 100.0% 85.9%
6 Bt_GP1_SLP91kDa_Coleoptera_AAY28601 100.0% 83.9%
7 Bt_B22_SLP-87kDa_ADU04488 100.0% 86.7%
8 Bt_CTC_SLP-EA1-91kDa_AAR23791 100.0% 87.4%
9 Bt_I13_SLP-EA1-91kDaADU04489 100.0% 86.4%
MAKTNSYKKVIAGTMTAAMVAGVVSPVAAAGKSFPDVPAGHWAEDSINYLVDKGAIVGKP
MAKTNSYKKVIAGTMTAAMVAGVVSPVAAAGKSFPDVPAGHWAEDSINYLVDKGAIVGKP
MAKTNSYKKVIAGTMTAAMVAGVVSPVAAAGKSFPDVPAGHWAEDSINYLVDKGAIVGKP
MAKTNSYKKVIAGTMTAAMVAGIVSPVAAAGKSFPDVPAGHWAEGSINYLVDKGAITGKP
MAKTNSYKKVIAGTMTAAMVAGVVSPVAAAGKSFPDVPAGHWLDSINYLVDKGAIEGKP
MAKTNSYKKVIAGTMTAAMVAGVVSPVAAAGKSFPDVPAGHWLDSINYLVDKGAIEGKP
MAKTNSYKKVIAGTMTAAMVAGVVSPVAAAGKTFPDVPAGHWLDSINYLVDKGAIEGKP
MAKTNSYKKVIAGTMTAAMVAGVVSPVAAAGKSFPDVPAGHWAEGSINYLVDKGAITGKP
MAKTNSYKKVIAGTMTAAMVAGVVSPVAAAGKSFPDVPAGHWAEGSINYLVDKGAITGKP
cov pid
1 Bt59S-SLP-EA1-91kDa_QDQ04690 100.0% 100.0%
2 Bt-soil-46-SLP-EA1-91kDa 100.0% 99.9%
3 Bt28s_SLP-EA1-91kDa 100.0% 99.9%
4 Bt_H67_SLP-EA1-91kDa_ADU04490 100.0% 90.6%
5 Bt_BMB1152_SLP-EA1-91kDa_ADU04491 100.0% 85.9%
6 Bt_GP1_SLP91kDa_Coleoptera_AAY28601 100.0% 83.9%
7 Bt_B22_SLP-87kDa_ADU04488 100.0% 86.7%
8 Bt_CTC_SLP-EA1-91kDa_AAR23791 100.0% 87.4%
9 Bt_I13_SLP-EA1-91kDaADU04489 100.0% 86.4%
DGTYGPTESIDRASAAVIFTKILNLPVDENAQPSFKDAKNTWAAKYIAAVEKAGIVKGDG
DGTYGPTESIDRASAAVIFTKILNLPVDENAQPSFKDAKNTWAAKYIAAVEKAGIVKGDG
DGTYGPTESIDRASAAVIFTKILNLPVDENAQPSFKDAKNTWAAKYIAAVEKAGIVKGDG
DGTYGPTESIDRASAAVIFTKILNLPVDENAQPSFKDGKNIWSSKYIAAVEKAGVVKGDG
DGTYAPAEDIDRASAAKIMAITLGLKVEEGAQPSFKDAKNHWASKYIAAVEKAGVVRGDG
DGTYAPAEEIDRASAAKIMAITLGLKVEEGAQPSFKDAKNHWASKYIAAVEKAGVVRGDG
DGTYAPAEEIDRASAAKIMAITLGLKVEEGAQPSFKDAKNHWASKYIAAVEKAGVVRGDG
DGTYGPTESIDRASAAVIFTKILNLPIDENAQPSFKDAKNIWSSKYIAAVEKAGVVKGDG
DGTYGPTESIDRASAAVIFTKILNLPIDENAQPSFKDAKNIWSSKYIAAVEKAGVVKGDG
cov pid
1 Bt59S-SLP-EA1-91kDa_QDQ04690 100.0% 100.0%
2 Bt-soil-46-SLP-EA1-91kDa 100.0% 99.9%
3 Bt28s_SLP-EA1-91kDa 100.0% 99.9%
4 Bt_H67_SLP-EA1-91kDa_ADU04490 100.0% 90.6%
5 Bt_BMB1152_SLP-EA1-91kDa_ADU04491 100.0% 85.9%
6 Bt_GP1_SLP91kDa_Coleoptera_AAY28601 100.0% 83.9%
7 Bt_B22_SLP-87kDa_ADU04488 100.0% 86.7%
8 Bt_CTC_SLP-EA1-91kDa_AAR23791 100.0% 87.4%
9 Bt_I13_SLP-EA1-91kDaADU04489 100.0% 86.4%
KENFYPEGKIDRASFASMLVSAYNLKDKVNGELVTTFDDLRKHWGEEKANILINLGISVG
KENFYPEGKIDRASFASMLVSAYNLKDKVNGELVTTFDDLRKHWGEEKANILINLGISVG
KENFYPEGKIDRASFASMLVSAYNLKDKVNGELVTTFDDLRKHWGEEKANILINLGISVG
KENFYPEGKIDRASFASMLVSAYNLKDKVNGELVTTFEDLLDHWGEEKANILINLGISVG
KENFSPDKKIDRASFASMIVGAYNLKDKVNGELVTKFEDLLDHWGEEKANILIHLGLSEG
KENFSPDKKIDRASFASMIVGAYNLKDKVNGELVTKFEDLLDHWGEEKANILIHLGLSEG
KENFSPDKRIDRASFASMIVGAYNLKDKVNGELVTKFEDLLDHWGEEKANILIHLGLSEG
KENFYPNGKIDRASFASMLVSAYNLKDKVSGELVTTFDDLRDTWGEEKANILINLGISVG
KENFYPNGKIDRASFASMLVSAYNLKDKVNGELVTTFEDLLDHWGEEKANILINLGISVG
cov pid
1 Bt59S-SLP-EA1-91kDa_QDQ04690 100.0% 100.0%
2 Bt-soil-46-SLP-EA1-91kDa 100.0% 99.9%
3 Bt28s_SLP-EA1-91kDa 100.0% 99.9%
4 Bt_H67_SLP-EA1-91kDa_ADU04490 100.0% 90.6%
5 Bt_BMB1152_SLP-EA1-91kDa_ADU04491 100.0% 85.9%
6 Bt_GP1_SLP91kDa_Coleoptera_AAY28601 100.0% 83.9%
7 Bt_B22_SLP-87kDa_ADU04488 100.0% 86.7%
8 Bt_CTC_SLP-EA1-91kDa_AAR23791 100.0% 87.4%
9 Bt_I13_SLP-EA1-91kDaADU04489 100.0% 86.4%
TGT-KWEPNKTVTRAEAAQFIALTDKKYGKKDNAQAYVTDVKISEPAKLTLTGTGLDKLS
TGT-KWEPNKTVTRAEAAQFIALTDKKYGKKDNAQAYVTDVKISEPAKLTLTGTGLDKLS
TGT-KWEPNKTVTRAEAAQFIALTDKKYGKKDNAQAYVTDVKISEPAKLTLTGTGLDKLS
TGG-KWEPNKSVSRAEAAQFIALTDKKYGKKDNAQAYVTDVKVSEPTKLTLTGTGLEKLS
TGGNKWEPDKSVSRAEAAQFIALTDKKYGKKDNAQAYVTDVKASEPTKLTLTGTGLDKLS
TGGNKWEPNKSVSRAEAAKFIAVADKKYGKKDNAQAYVTDVKVSEPTKLTLTGTGLDKLA
TGGNKWEPNKSVSRAEAAQFIAVADKKYGKKDNAQAYVTDVKVSEPTKLTLTGTGLDKLA
TGG-KWEPTKSVSRAEAAQFIALTDKKYGKKDNAQAYVTDVKVSEPAKLTLTGTGLDKLS
IDG-KWEPNKTVSRAEAAQFIALTDKKYGKKDNAQAYVTDVKVSEPAKLTLTGTGLDKLS
cov pid
1 Bt59S-SLP-EA1-91kDa_QDQ04690 100.0% 100.0%
2 Bt-soil-46-SLP-EA1-91kDa 100.0% 99.9%
3 Bt28s_SLP-EA1-91kDa 100.0% 99.9%
4 Bt_H67_SLP-EA1-91kDa_ADU04490 100.0% 90.6%
5 Bt_BMB1152_SLP-EA1-91kDa_ADU04491 100.0% 85.9%
6 Bt_GP1_SLP91kDa_Coleoptera_AAY28601 100.0% 83.9%
7 Bt_B22_SLP-87kDa_ADU04488 100.0% 86.7%
8 Bt_CTC_SLP-EA1-91kDa_AAR23791 100.0% 87.4%
9 Bt_I13_SLP-EA1-91kDaADU04489 100.0% 86.4%
ADDVTLEGDKAVAIEASADGTSAVVTLGGKVAPNKDLSIKVKNQSFLAKFVYEVKKLAVE
ADDVTLEGDKAVAIEASADGTSAVVTLGGKVAPNKDLSIKVKNQSFVAKFVYEVKKLAVE
ADDVTLEGDKAVAIEASADGTSAVVTLGGKVAPNKDLSIKVKNQSFVAKFVYEVKKLAVE
ADDVTLEGDKAVAIEASADGTSATVTLGGKVAPNKNLTVKVKNQSFVTKFVYEVKKLEVE
ADDVTLEGDKAVAIEASTDGTSAVVTLGGKVAPNKDLTVKVKNQSFVTKFVYEVKKLAVE
AEDVTLEGDKAVAIEASADGTSAVVTLGGKVAPNKNLTVKVKNQSFVTKFVYEVKKLAVE
ADDVTLEGDKAVAIEASADGTSAVVTLGGKVAPNKNLTVKAKNQSFVTKFVYEVKKLAVE
ADDVTLEGDKAVAIEASADGNSATVTLGGKVAPNKNLTVKVKNQSFVTKFVYEVKKLAVE
ADDVTLEGDKAVAIEASADGNSATVTLGGKVAPNKNLTVKVKNQSFVTKFVYEVKKLAVE
cov pid
1 Bt59S-SLP-EA1-91kDa_QDQ04690 100.0% 100.0%
2 Bt-soil-46-SLP-EA1-91kDa 100.0% 99.9%
3 Bt28s_SLP-EA1-91kDa 100.0% 99.9%
4 Bt_H67_SLP-EA1-91kDa_ADU04490 100.0% 90.6%
5 Bt_BMB1152_SLP-EA1-91kDa_ADU04491 100.0% 85.9%
6 Bt_GP1_SLP91kDa_Coleoptera_AAY28601 100.0% 83.9%
7 Bt_B22_SLP-87kDa_ADU04488 100.0% 86.7%
8 Bt_CTC_SLP-EA1-91kDa_AAR23791 100.0% 87.4%
9 Bt_I13_SLP-EA1-91kDaADU04489 100.0% 86.4%
KLTFDDDRAGQAIVFKLNDEKGNADVEYLNLANHDVKFVANNLDGSPANIFEG----GKS
KLTFDDDRAGQAIVFKLNDEKGNADVEYLNLANHDVKFVANNLDGSPANIFEG----GKS
KLTFDDDRAGQAIVFKLNDEKGNADVEYLNLANHDVKFVANNLDGSPANIFEG----GKS
KLTFDDDRAGQAVVFKLNDEKGNADVEYLNLADHDVKFVANNLDGSPAYIFEN----GEA
KLTFDDDRAGQAIAFKLNDEKGNADVEYLNLANHDVKFVANNLDGSPANIFEG----GEA
KLTFDDDRAGQAVAFKLNDEKGNADVEYLNLADHDVKFVANNLDGSSANIFEG----GVA
KLTFDDDRAGQSVVFKLNDEKGNADVEYLNLADHDVKFVANNLDGSPAYIFEN----GEA
KLTFDDDRAGQAVAFKLNDEKGNADVEYLNLADHDVKFVANNLDGSSANIFEG----GVS
KLTFDDDRAGQAIAFKLNDEKGNADVEYLNLADHDVKFVANKLDGSPAKIFEGKDGKKEA
cov pid
1 Bt59S-SLP-EA1-91kDa_QDQ04690 100.0% 100.0%
2 Bt-soil-46-SLP-EA1-91kDa 100.0% 99.9%
3 Bt28s_SLP-EA1-91kDa 100.0% 99.9%
4 Bt_H67_SLP-EA1-91kDa_ADU04490 100.0% 90.6%
5 Bt_BMB1152_SLP-EA1-91kDa_ADU04491 100.0% 85.9%
6 Bt_GP1_SLP91kDa_Coleoptera_AAY28601 100.0% 83.9%
7 Bt_B22_SLP-87kDa_ADU04488 100.0% 86.7%
8 Bt_CTC_SLP-EA1-91kDa_AAR23791 100.0% 87.4%
9 Bt_I13_SLP-EA1-91kDaADU04489 100.0% 86.4%
TSSTGKLGVGLSHGDYKVEVQVTKRGGLTVSNTGIISVKNLDTPASAIKSTVFAVDADKN
TSSTGKLGVGLSHGDYKVEVQVTKRGGLTVSNTGIISVKNLDTPASAIKSTVFAVDADKN
TSSTGKLGVGLSHGDYKVEVQVTKRGGLTVSNTGIISVKNLDTPASAIKSTVFAVDADKN
TSATGKLGVGLSHGDYKVEVQVTKRGGLTVSNTGIISVKNLDTPASAIKSTAFAVDADKN
TSTTGKLAVGIKQGDYKVEVQVTKRGGLTVSNTGIITVKNLDTPASAIKNVFALDADND
TSTTGKLAVGIKPADYKVEVQVTKRGGLTVSNTGIITVKNLDTPASAIKNAVFALDADND
TSSTGKLGVGLSHGDYKVEVQVTKRGGLTVSNTGIISVKNLDTPASAIKNTVFAVDADKN
TSTTGKLAVGITPADYKVEVQVTKRGGLTVSNTGIITVKNLDTPASAIKNVFALDADND
TSATGKLAVGITPADYKVEVQVTKRGGLTVSNTGIITVKNLDTPASAIKNVFAVDADSD

Continued

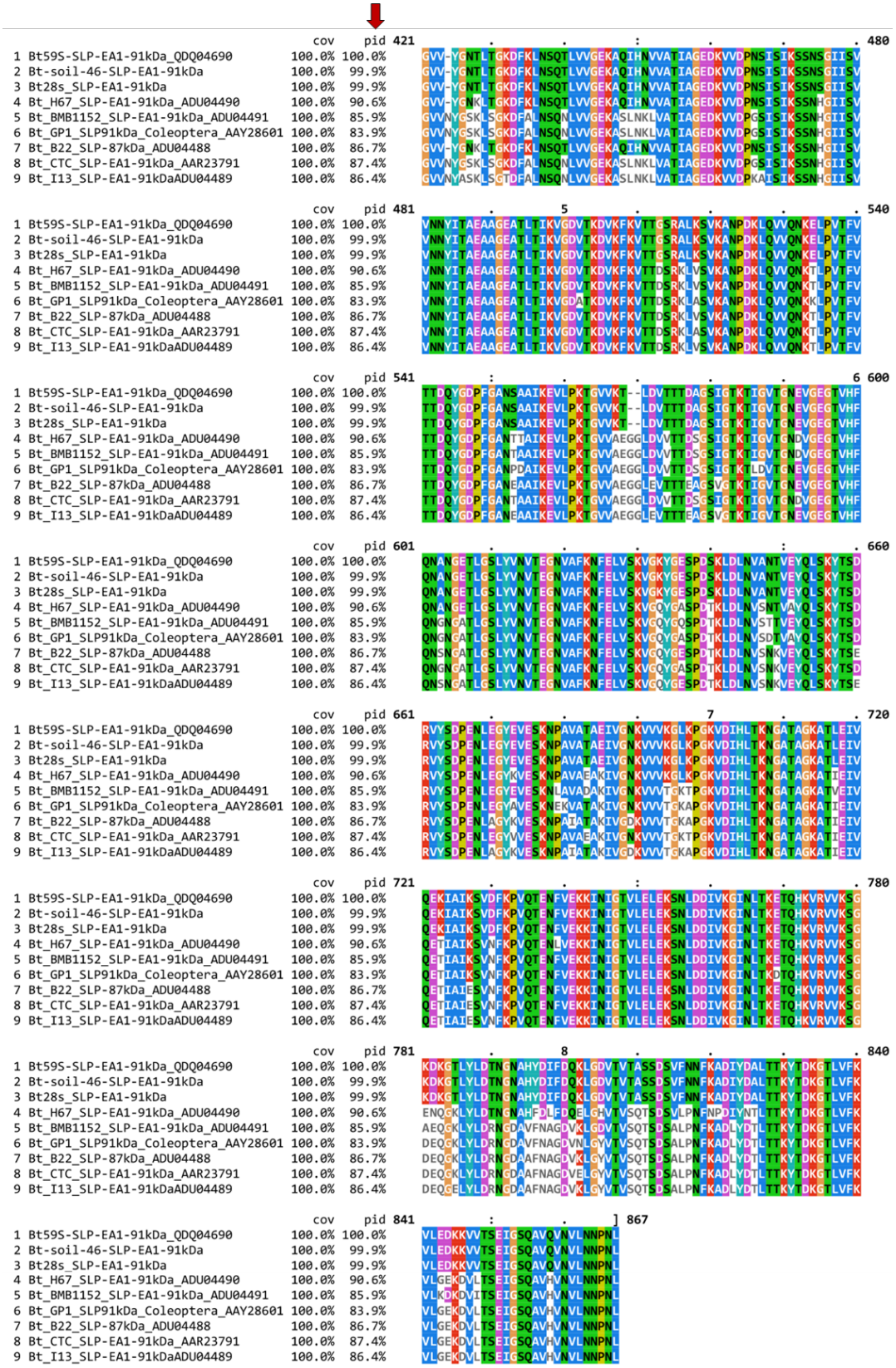
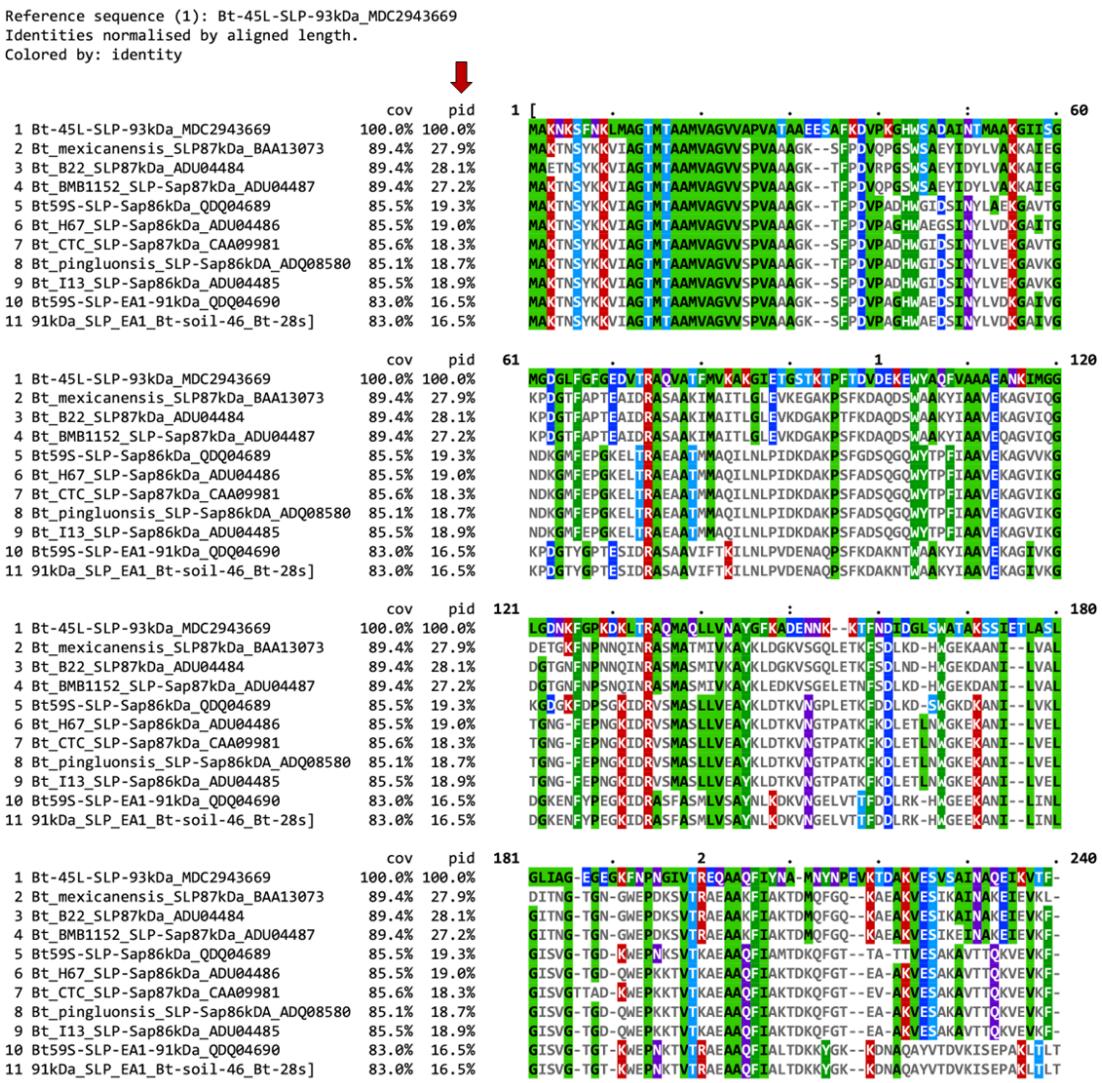



**Figure 6.3.11:** MSA and percent identity of 91 kDa EA1 of *B. thuringiensis* strains BD59S, Soil-46, and 28S with parasporal inclusion forming EA1 of other Bt strains. (pid means percent identity)

***MSA and percent identity of 93 kDa SLP of 45L with indigenous and closely parasporal inclusion forming SLP of other Bt strains:***

The 93 kDa S-layer protein of the indigenous Bt genome 45L had 27% to 18% sequence identity with 7 parasporal SLP and 19% to 16% sequence identity with indigenous Sap and EA-1. The 93 kDa S-layer protein of indigenous 45L had 27–28% sequence similarities with a special type of 87 kDa SLP of three Bt strains: B22, mexicanensis, and BMB1152 (Figure 6.3.12). As parasporal SLPs are genealogically heterogeneous, so, based on the MSA and percent identity, it seems likely that the 93 kDa SLP of indigenous Bt strain 45L and the parasporal inclusion formation SLP of other Bt strains are related. So, it could be parasporal inclusion-forming protein of 45L.




Continued

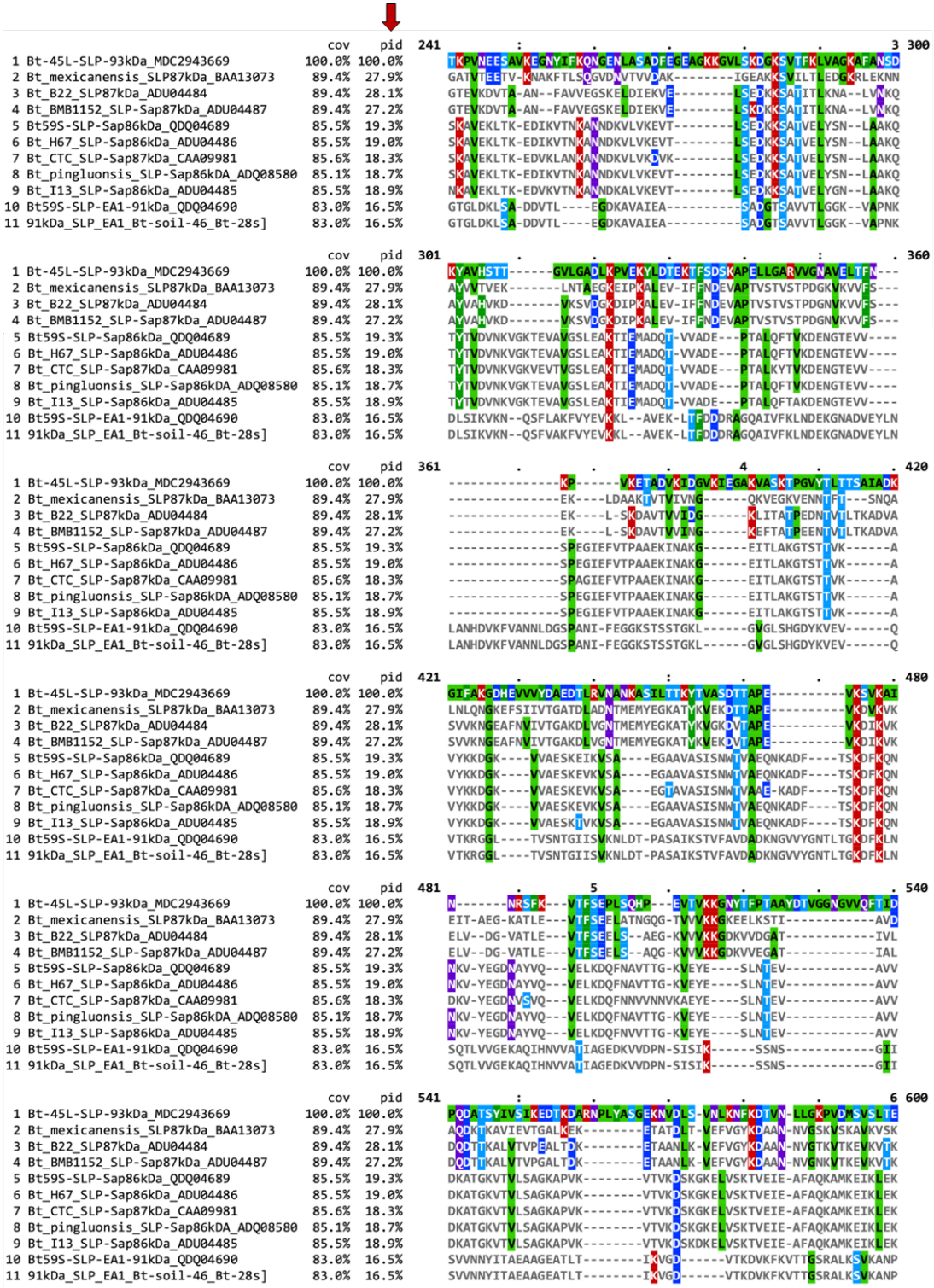



**Figure 6.3.12:** MSA and percent identity of 93 kDa SLP of 45L with indigenous and closely related parasporal inclusion forming SLP of other Bt strains.

***MSA and percent identity of 103 kDa SLP of 45L with indigenous and closely related parasporal inclusion forming SLP of other Bt strains:***

The 103 kDa SLP of indigenous Bt strain 45L had 16% sequence similarities with 6 parasporal SLPs of different Bt strains and it had maximum sequence similarities with EA-1 (Figure 6.3.12). Besides, the 103 kDa SLP had 8% sequence similarities with the 93 kDa SLP of its own strain and 12% sequence similarities with the Sap of indigenous 3 Bt strains. Although the overall percent identity was very poor with the other parasporal SLP, but it had significant similarities with SLH region (Figure 6.3.13). Besides, based on the similarities of the size of the 103 kDa SLP and the SDS-PAGE observed ~103 kDa parasporal protein band, it can be assumed that the 103 kDa SLP of the 45L could be the 103 kDa parasporal inclusion forming SLP of this strain.

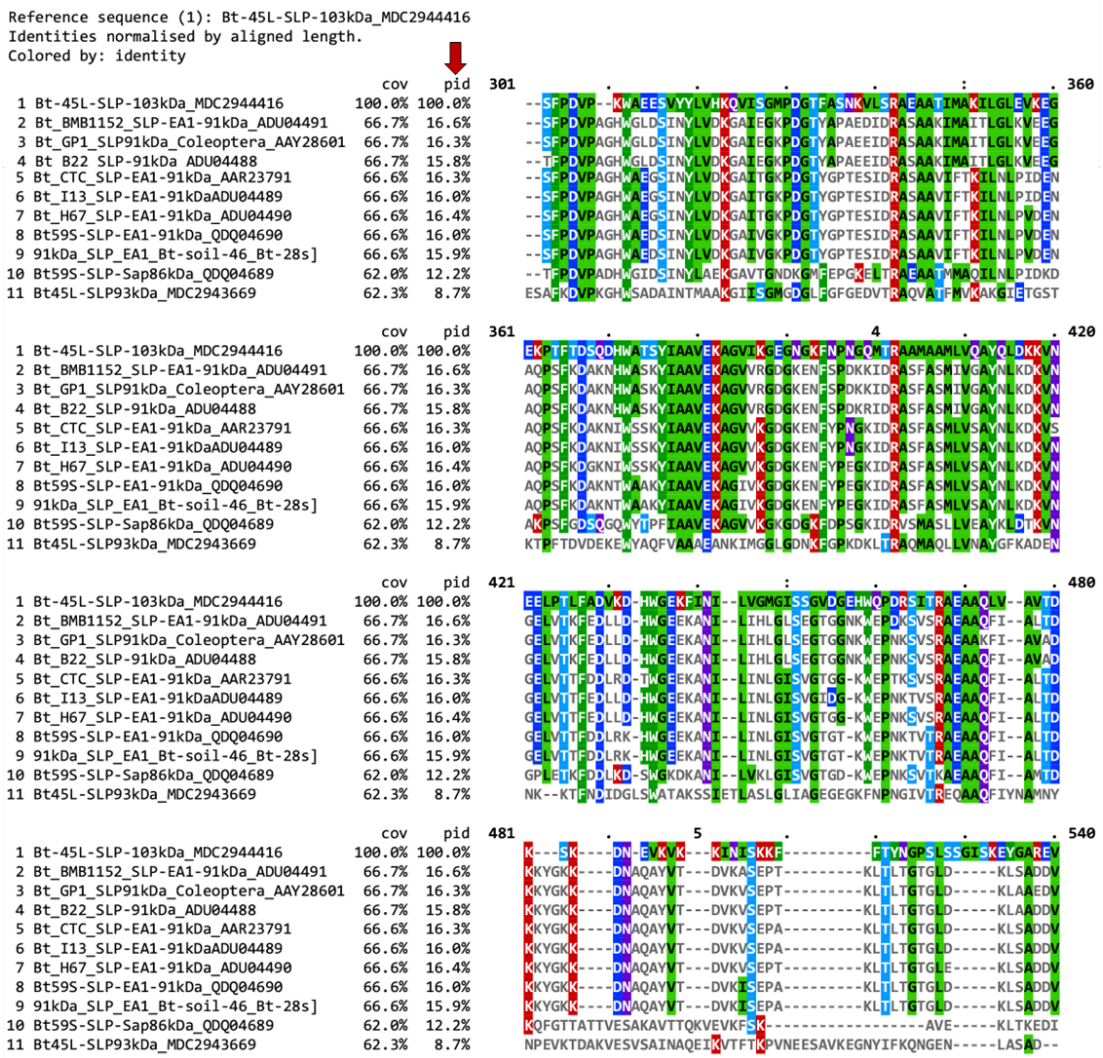



**Figure 6.3.13:** MSA and percent identity of 103 kDa S-layer protein of *B. thuringiensis* strain 45L with indigenous and closely related parasporal inclusion forming SLP of other Bt strains.

#### 6.3.4.3 Phylogenetic relationship of indigenous SLP with parasporal inclusion forming S-layer protein

A phylogenetic tree was constructed using the amino acid sequences of indigenous SLPs and the 13 parasporal inclusions forming SLPs of other Bt strains. The phylogenetic tree was split into two major nodes. Each node was again divided into different clades and taxa (Figure 6.3.14). The phylogenetic tree showed that the 86 kDa Sap and the 91 kDa EA1 were located in two separate nodes.

The 103 kDa SLP of the indigenous 45L was split from the 91 kDa EA1 at the first node, which was found in the upper part of the tree, and formed a single isolated clade (Figure 6.3.14). The 91-kDa EA1s formed a single clade and were split into two different taxa. The taxa and clades in the 91 kDa EA1 of the tree made it clear that all of the native 91 kDa EA1s were closely linked to the parasporal inclusion forming EA1 of other Bt. Furthermore, while the 103 kDa SLP of 45L formed a separate clade and taxon within this node, it was thought to connect the EA1 group of SLPs.

On the other hand, in the second or lower positioning node of the tree, the 93 kDa SLP of indigenous 45L strains and the 87 kDa parasporal inclusion forming 3 SLPs of other countries of origin produced a single clade and they were different from the 86 kDa SLPs Sap (Figure 6.3.14). The 86 kDa Sap of indigenous 3 Bt strains and 4 other 86 kDa Sap created a single clade; they were again divided into two separate taxa. From this section of the phylogenetic tree, it was clear that indigenous 86 kDa Saps were very close to the parasporal inclusion forming 86 kDa Saps of other countries of origin. Besides, it was also clear that the 93 kDa SLP of the 45L had close similarities with the parasporal inclusion forming 87 kDa SLPs as well as the 86 kDa SLPs.

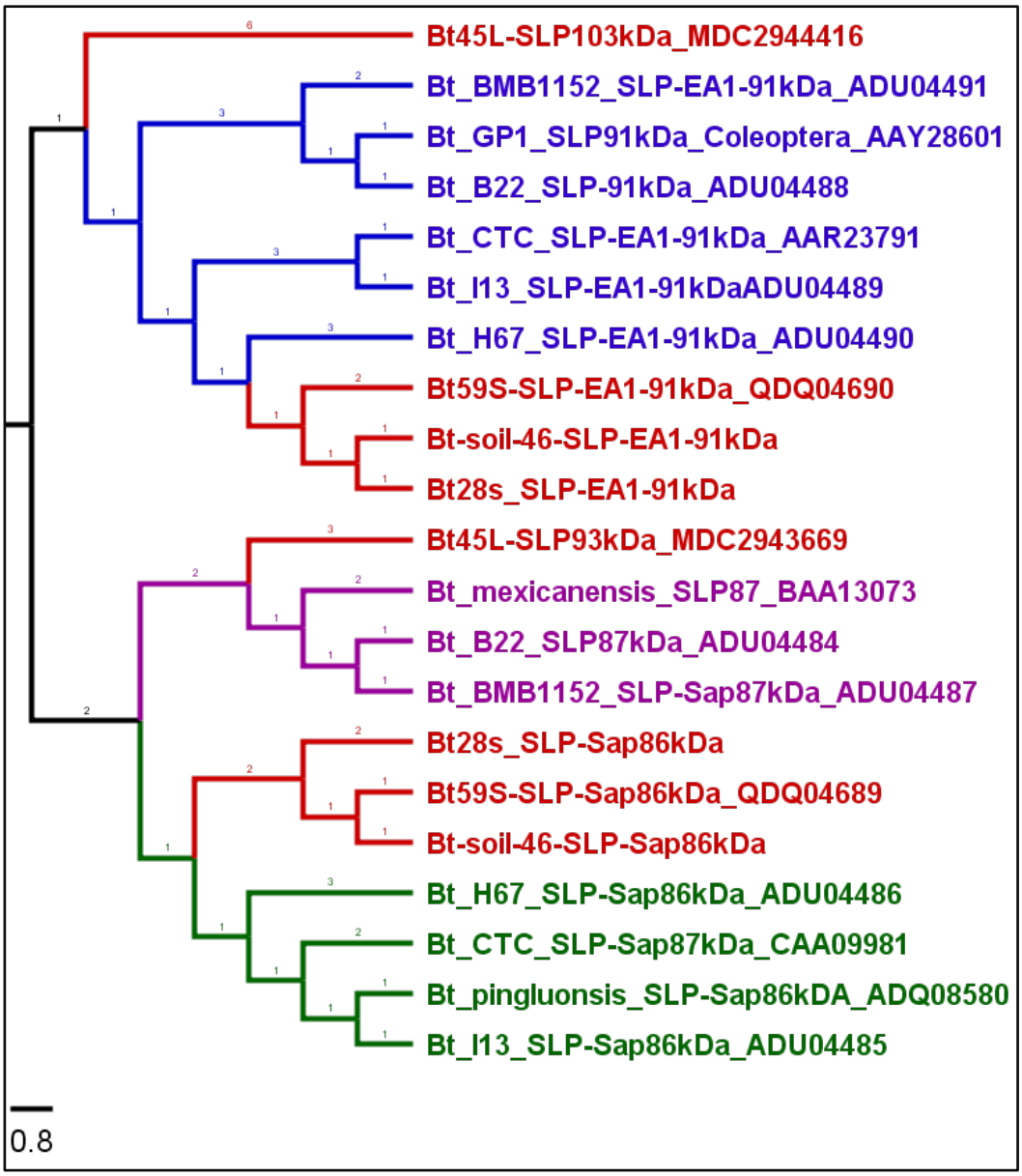



**Figure 6.3.14:** Phylogenetic tree showing the relationships between indigenous and parasporal inclusion forming SLP of other Bt strains. (Accession number is at the end of each taxon).

### 6.3.4.4 Genetic organization of *sap* and *eag* in the genome

This study looked at how the genetic features of the surface array protein (Sap) synthesizing gene *sap* and the extractable antigen 1 (EA1) protein synthesizing gene *eag* were organized. It used the BV-BRC feature search and comparison tool. The *sap* and *eag* genes were clearly found in the genomes of native *B. thuringiensis* strains BD59S, 28S, and Soil-46. They were also found in tandem pattern in the genomes of

*B. thuringiensis, B. anthracis*, and *B. cereus* (Figure 6.3.15). This result supported the previously published findings of Soufiane *et al.*, (2011), where they found that the *sap* and *eag* genes in *B. thuringiensis* were either present in tandem in 35% or absent in 65% of the strains.

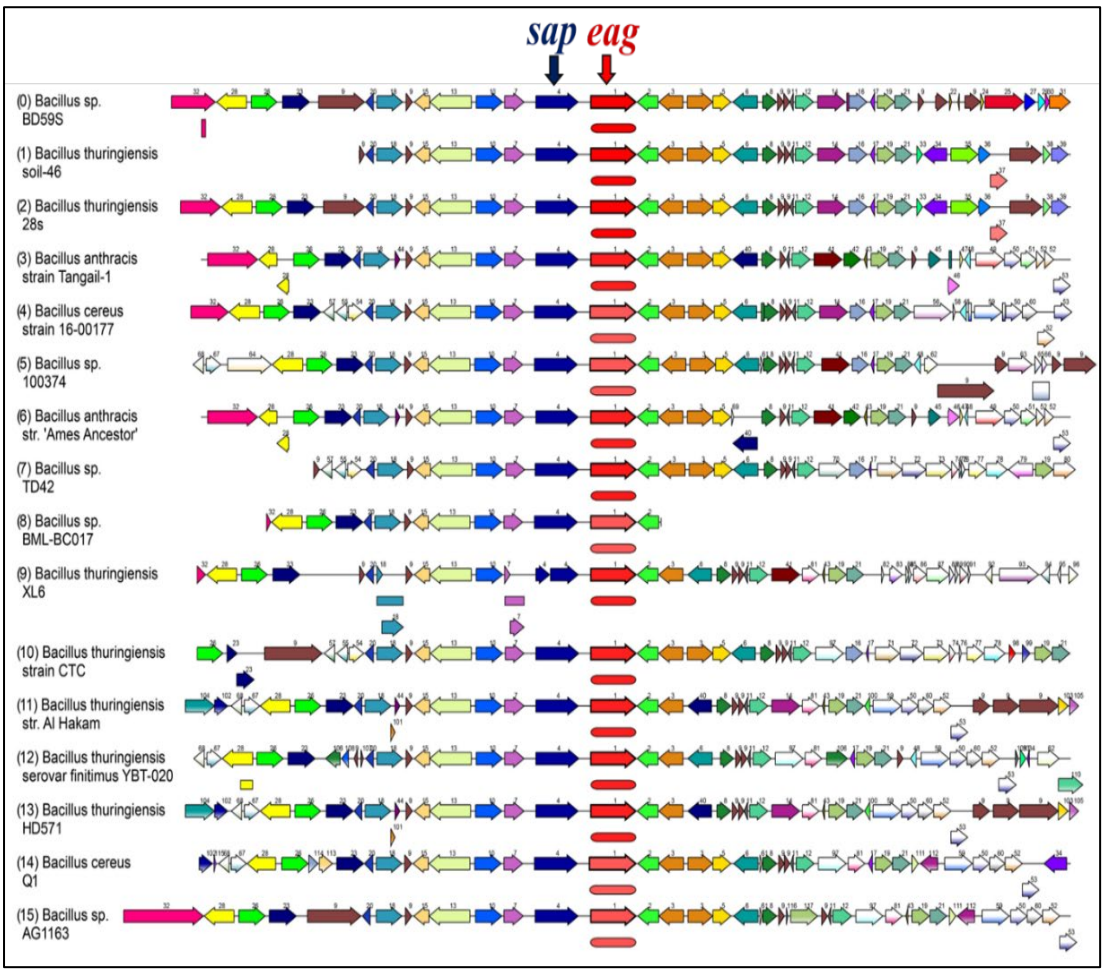



**Figure 6.3.15:** Genetic organization of the *sap* and *eag* genes in the genome.

We also looked at how the genetic features of S-layer genes in native *B. thuringiensis* strain 45L were organized. It was also studied by BV-BRC feature search and comparison tools. The 93 kDa and 103 kDa protein synthesizing genes were not found closely situated. Instead, they were found in two different places in the genome of *B. thuringiensis* strain 45L and other bacteria in the *B. thuringiensis* and *B. cereus* groups (Figures 6.3.16 and 6.3.17).

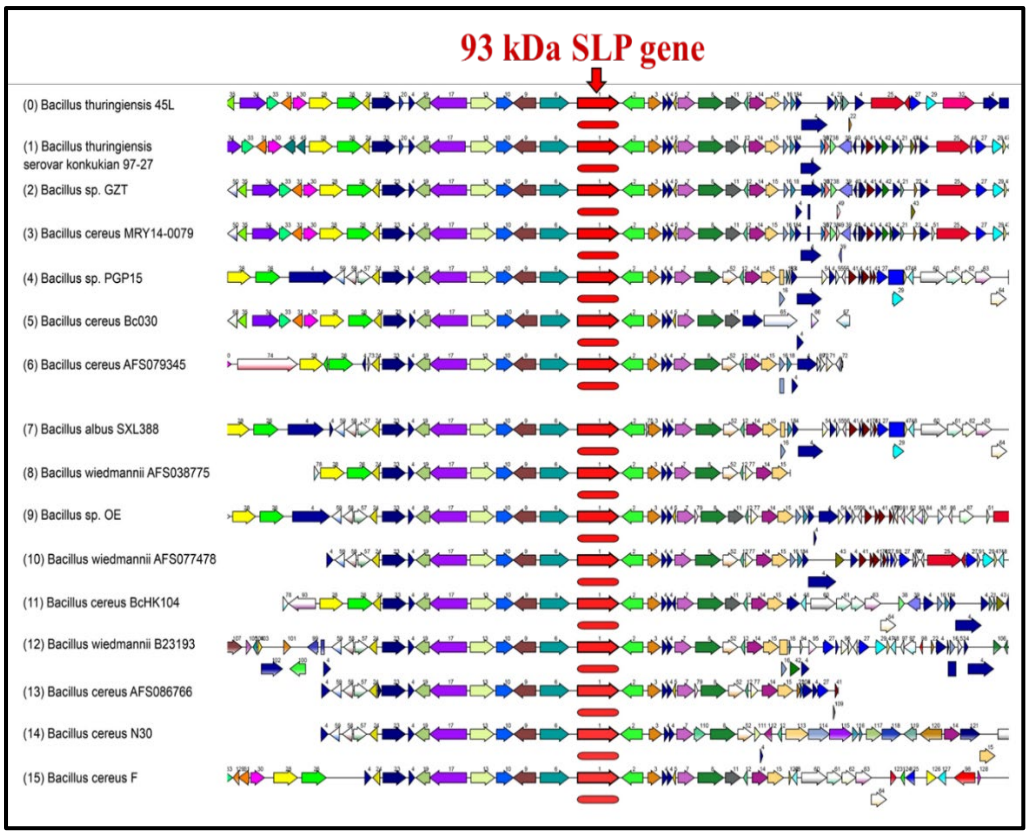



**Figure 6.3.16:** Genetic organization of the 93 kDa S-layer gene in the genome.

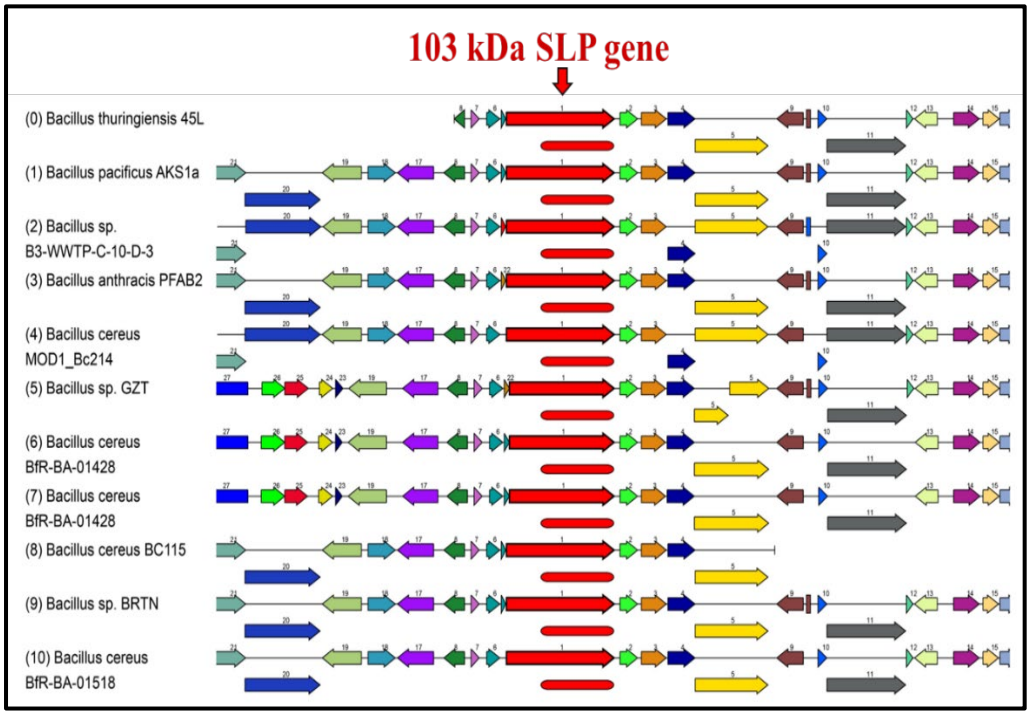



**Figure 6.3.17:** Genetic organization of the 103 kDa S-layer gene in the genome.

#### 6.3.4.5 Comparative motif study of indigenous and parasporal inclusion forming S-layer proteins of Bt strains.

The full-length S-layer protein sequences were used to execute the motif search on the MEME online tool. A total of 21 S-layer proteins were studied for motif analysis; of them, 13 were parasporal inclusion-forming and 8 were indigenous S-layer proteins. Through the Multiple Expectation maximizations for Motif Elicitation (MEME) analysis, a total 10 conserved motifs were found from 21 S-layer proteins. These were shown in different colored boxes. Based on motif architecture, they were divided into 5 different groups (Figure 6.3.18). Of group 1, the 103 kDa SLP of 45L was found to have only 3 conserved motifs (Motif 2-4). These three conserved motifs (Motif 2-3) were found common in all of the 21 S-layer proteins. Besides, nine S-layer proteins create a separate group (group 2) and each of them contains 10 conserved motifs. The 91 kDa EA1 from three native Bt strains (BD59S, Soil-46, and 28S) and the 91 kDa parasporal inclusion forming EA1 from six Bt strains (BMB1152, GP1, CTC, IT3, H67, and B22) formed group 2. These strains all had the identical 10 conserved motifs. Besides, group 3, the 93 kDa SLP of 45L was found to have only 4 conserved motifs (Motif 1-4). Group 4 had 3 S-layer proteins and was found to have 5 conserved motifs (Motif 1-4 and Motif 10). The 87 kDa SLP of Bt strains mexicanensis, B22, and BMB1152 made group 4. The 86 kDa Sap of 3 indigenous Bt strains BD59S, Soil-46, and 28S and the 86 kDa parasporal inclusion forming Sap of 4 Bt strains H67, CTC, pingluonsis, and IT3 made an identical structural cluster of 9 motifs (Motif 1-6 and Motif 8-10) and formed group 5.

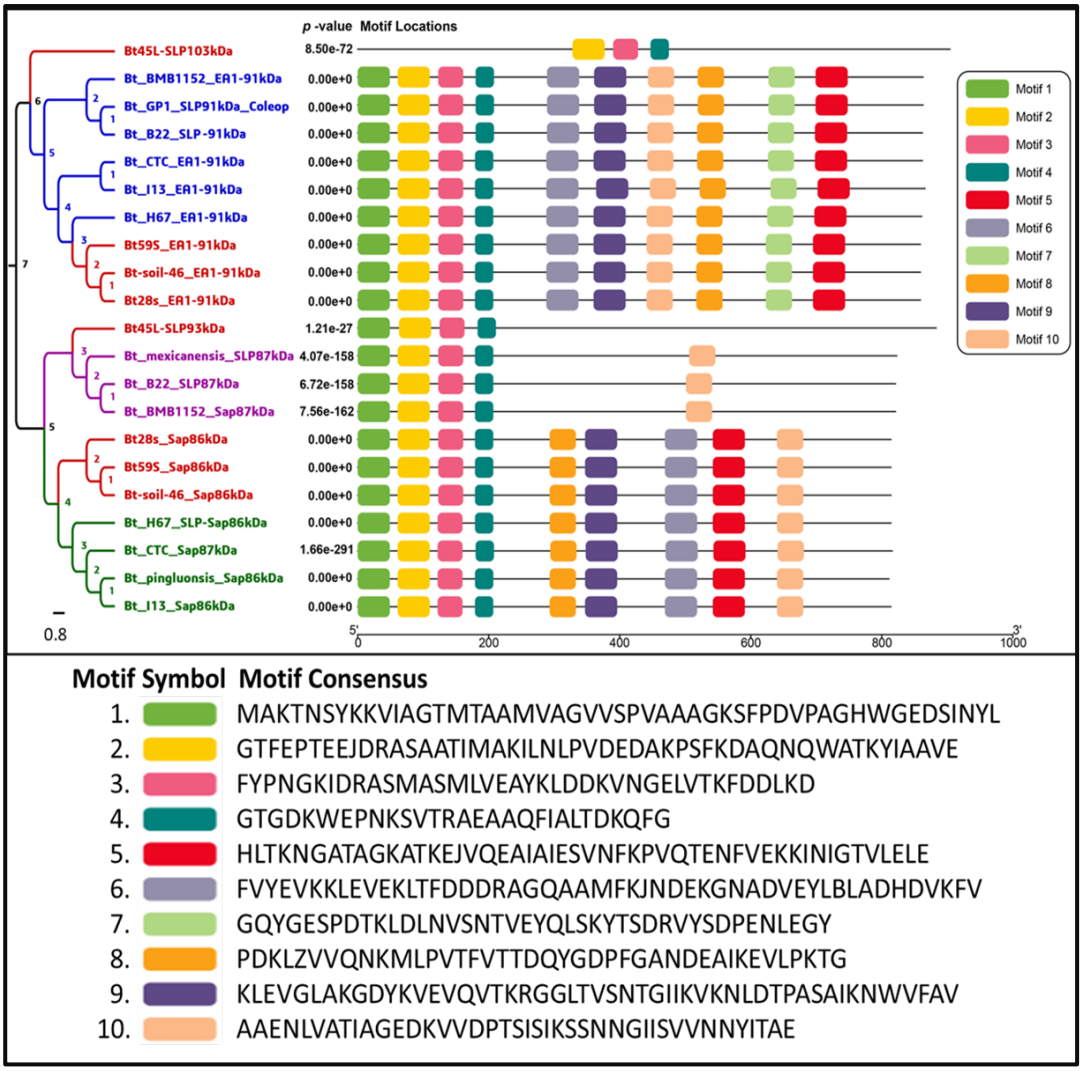



**Figure 6.3.18:** Conserved motif architecture of indigenous and parasporal SLP of Bt strains. Motifs were indicated in different colored boxes.

### 6.3.4.6 Comparative domain study of indigenous and parasporal inclusion forming S-layer proteins of Bt strains.

The full-length S-layer protein sequences were used to find out the conserved domains from the NCBI CD-search tool and the resulting data were viewed by the TB tool. A total of 21 S-layer proteins were studied for domain analysis; of them, 13 were parasporal inclusion-forming and 8 were indigenous S-layer proteins. Through the NCBI CD-search, three surface layer homology (SLH) conserved domains were found in all 21 S-layer proteins. These were shown in red boxes in Figure 6.3.19. Besides, the 3 SLH domains, 3 different types of additional domains, TolA, YvpB, and SH3, were also found in the 103 kDa SLP of Bt 45L. The 86 kDa saps of three native Bt strains

BD59S, Soil-46, and 28S also had two other types of domains: Big and BID. These were also found in the parasporal inclusion-forming Saps of two Bt strains H67 and pingluonsis. In addition, the 86 kDa parasporal inclusion forming Sap of Bt strain CTC and I13 also had a Big. The 86-kDa Sap of CTC also had a PRK08026 domain.

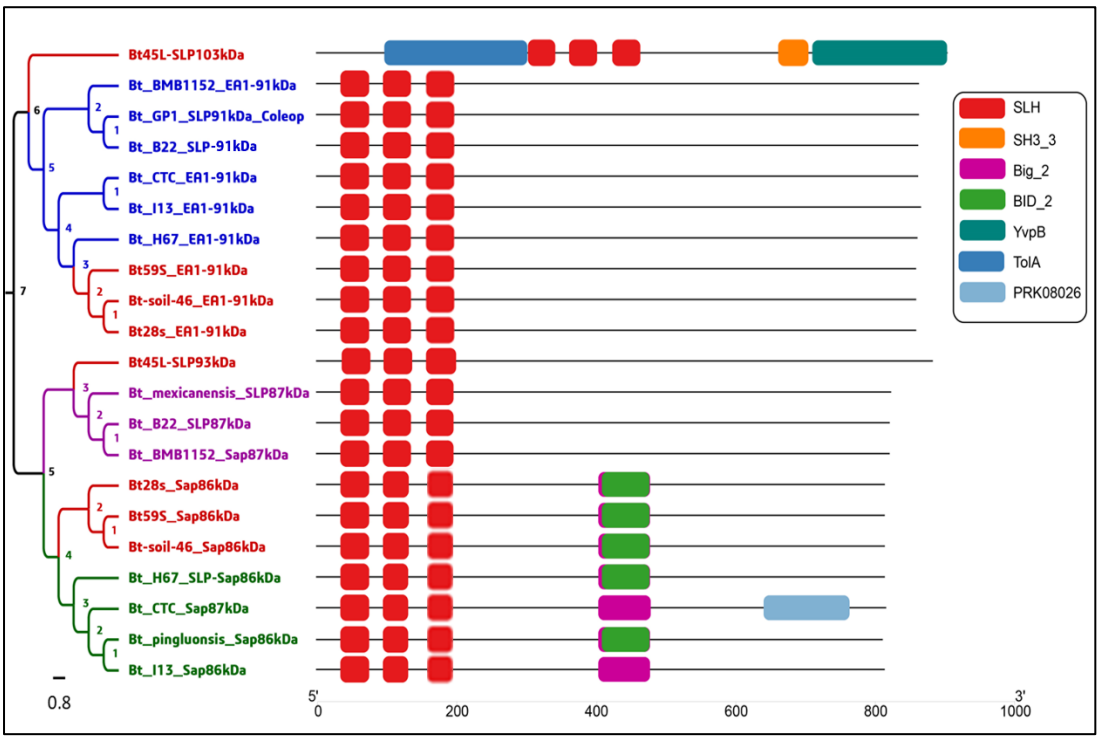



**Figure 6.3.19:** Conserved domain architecture of indigenous and parasporal SLP of Bt strains.

### 6.3.4.7 Functional domain and structure study of the S-layer and Cry-like proteins of indigenous Bt strains

#### 6.3.4.7.1 Domain study of the 86 kDa Sap of BD59S, Soil-46, and 28S

From the InterPro scan, NCBI protein features study, NCBI conserved domain search (CDD), and simple modular architecture research tool (SMART), it was found that the 86 kDa S-layer protein Sap of indigenous Bt strains BD59S, Soil-46, and 28S had a signal peptide (amino acids 1 to 30) and 3 S-layer homology domains (SLH) (from amino acids 33-76, 94-136, and 156-197) (Figure 6.3.20). Besides, from *in silico* tertiary and secondary protein structure prediction and domain architecture observation, it was found that each of the 3SLH domains contained α-helices and loops (Figure 6.3.21. A-B). In addition, β-strand-rich six β-sandwich crystallization or assembly

domains (from 215 to 805 amino acids) were also found in the tertiary and secondary structures of SLP Sap (Figure 6.3.21. A, C). Besides, 2 bacterial Ig-like domains (Big2) (from amino acids 402 to 481, and 510 to 579) were identified from the SMART search tool, and in a tertiary and secondary structure study, it was found that they were located in the β-strand-rich major protein portion (Figure 6.3.21. A). Besides, from the InterPro scan, an immunoglobulin-like copper resistance region CopC/internalin (213-304) and an invasin/intimin cell adhesion region (417-476) were also found in the 86 kDa Sap.

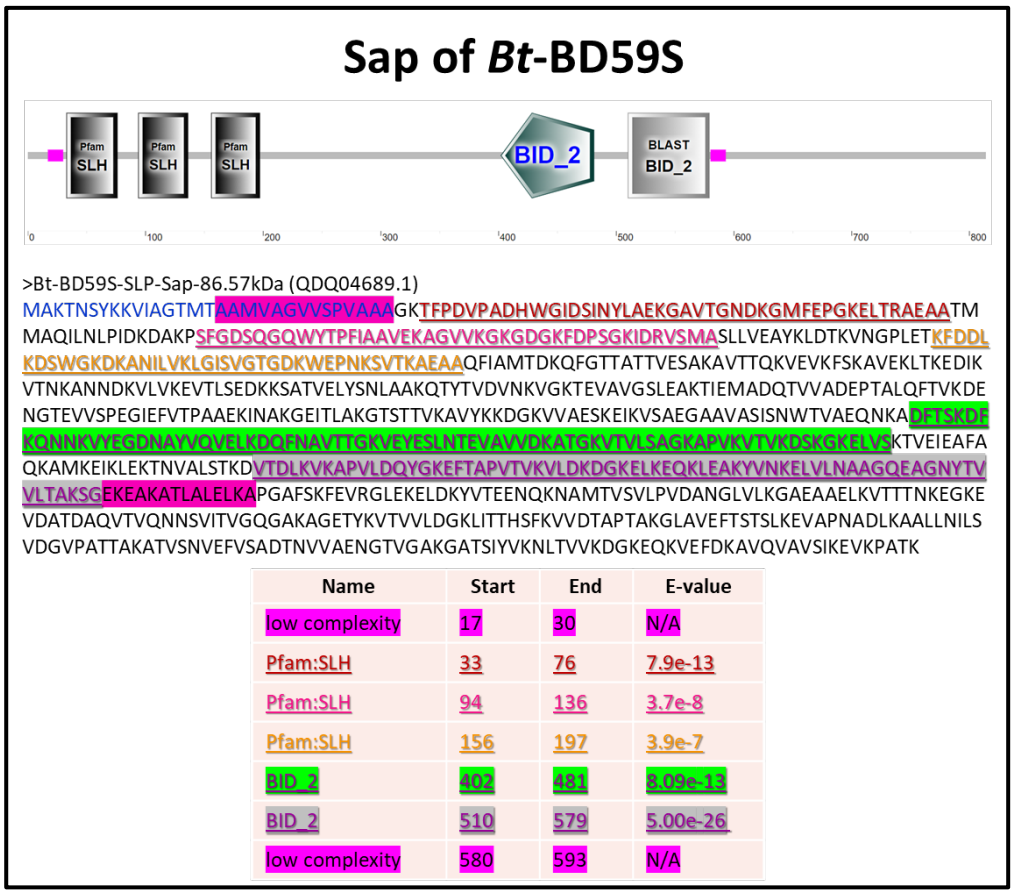



**Figure 6.3.20:** Schematic diagram of domains and features of the Sap of BD59S, Soil-46, and 28S.

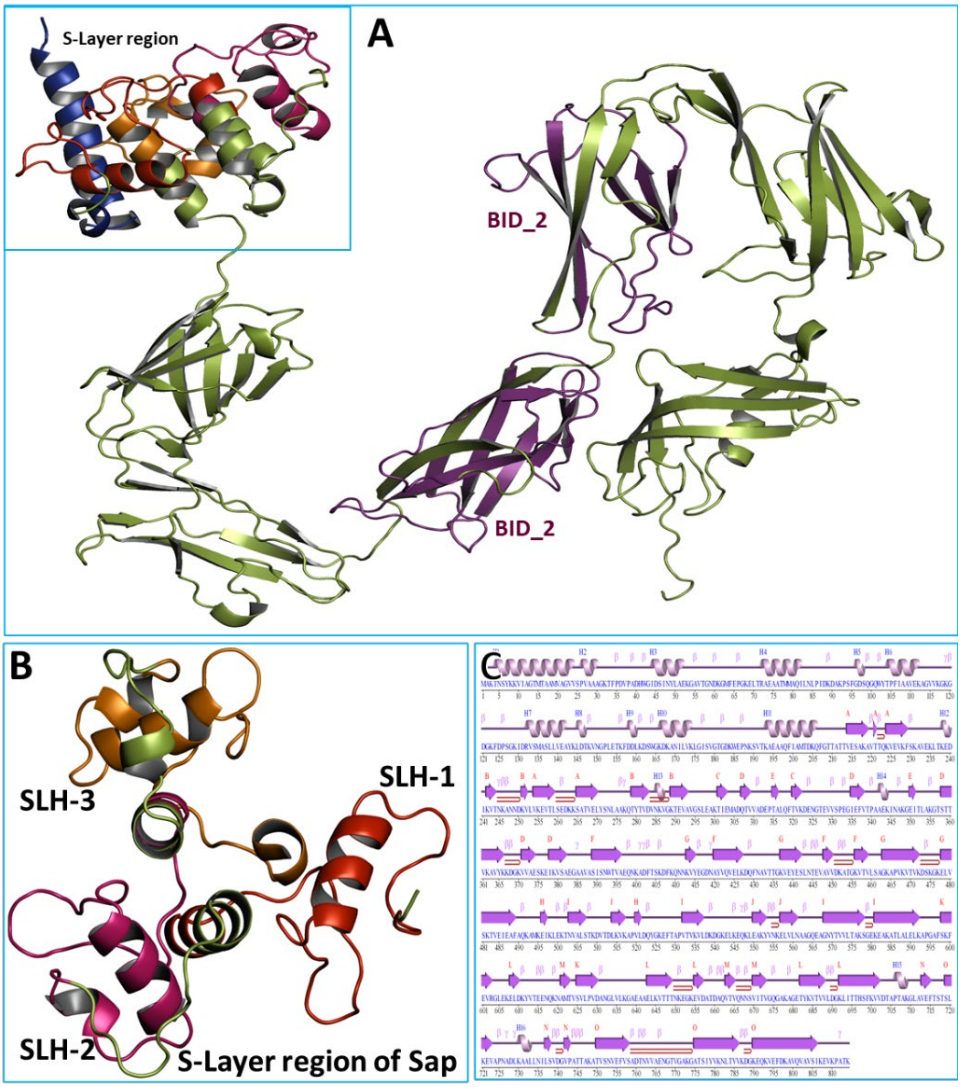



**Figure 6.3.21:** 3D and 2D protein structures of Sap of BD59S, Soil-46, and 28S. (A) 3D tertiary protein structure with SLH domains and β-strand-rich crystallization domains; (B) close view of 3 SLH domains; and (C) secondary protein structure of 91 kDa S-layer protein Sap

#### 6.3.4.7.2 Domain study of the 91 kDa EA1 of BD59S, Soil-46, and 28S

From the protein domain study, it was found that the 91 kDa EA1 had a signal peptide (amino acids 1-29) and 3 S-layer homology domains (SLH), of which domain I was from amino acids 33 to 76, domain II was from amino acids 94 to 136, and domain III was from amino acids 156 to 197 (Figure 6.3.22). Besides, from *in silico* tertiary and secondary protein structure prediction and domain architecture observation, it was found that each of the SLH domain contained α-helices and loops (Figure 6.3.23.A-B). Besides, β-strand-rich six β-sandwich crystallization or assembly domains (from 218 to 852 amino acids) were also found in the tertiary and secondary structures of the EA1 (Figure 6.3.23. A, C).

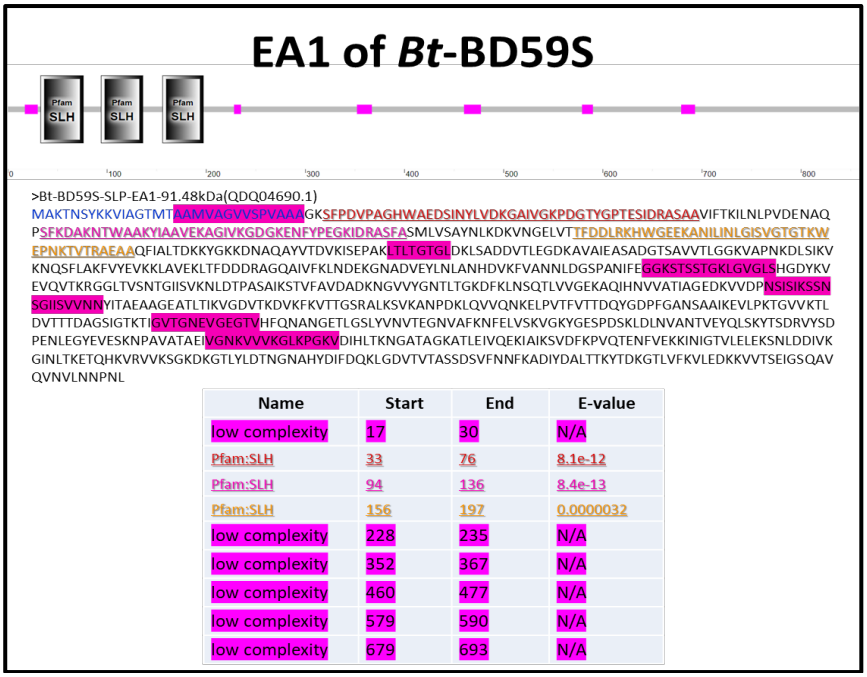



**Figure 6.3.22:** Schematic diagram of domains and features of the EA of BD59S, Soil-46, and 28S.

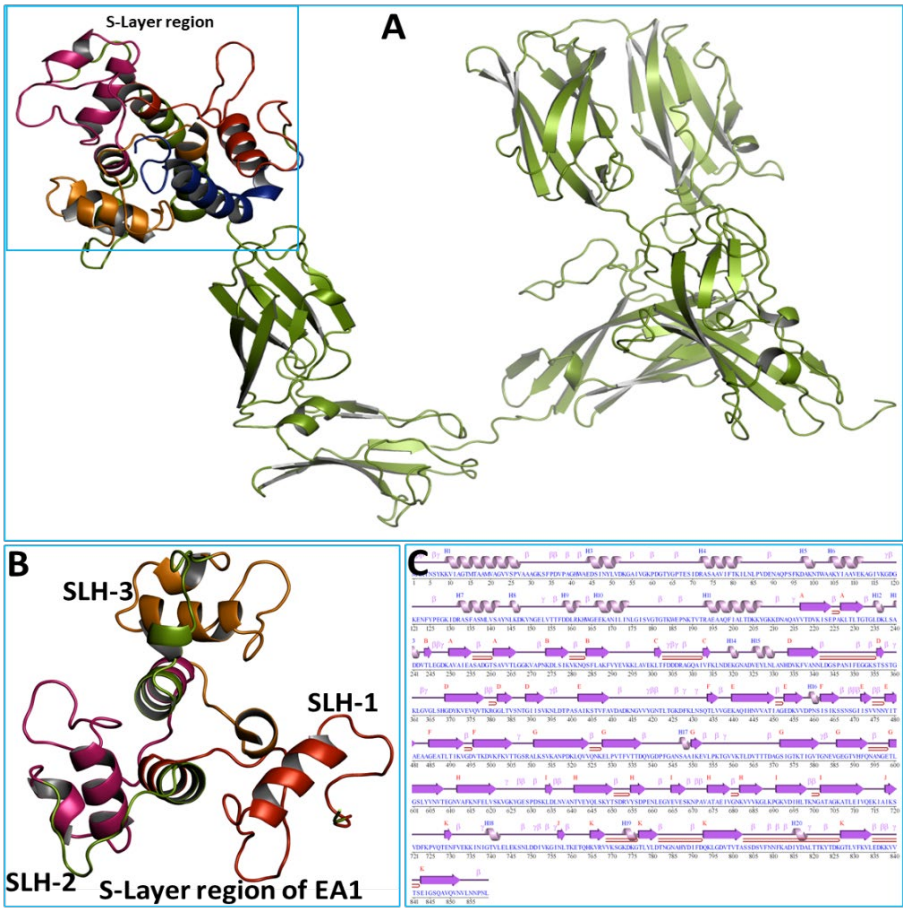



**Figure 6.3.23:** 3D and 2D protein structures of EA1 in indigenous BD59S. (A) Tertiary protein structure with SLH domains and β-strand-rich crystallization domains; (B) close view of 3 SLH domains; and (C) the secondary protein structure of 91 kDa S-layer protein EA1.

Besides, a single amino acid change was found in the protein sequences of EA1 of BD59S. It was also found that trypsin-digested parasporal protein fragments of BD59S were different from those of Soil-46 and 28S because they didn't produce 45 kDa and 32 kDa trypsin-digested protein fragments in BD59S, whereas Soil-46 and 28S did; hence, bioinformatical analysis was conducted based on nucleotide alignment, protein's secondary structure comparison and 3D pdb structure alignment. As a result of single nucleotide polymorphism (Figure 6.3.24. A), the single amino acid was varied (Figure 6.3.24. B-C), and their distinct changes were observed both in the secondary and tertiary structures (Figure 6.3.24. B-C). From the secondary structure, it was found that EA1 of BD59S had a valine in 286 positions whereas a leucine was present in EA1 of Soil-46 and 28S. As a result, EA1 of BD59S had an alpha helix in 320 amino acid positions but instead an alpha helix in EA1 of Soil-46 and 28S had a loop (Figure 6.3.24. B). From the 3D tertiary structure alignment between these two EA1, it was clear that the position of the first β-strand-rich domain (after the 3 SLH domain) differentiated easily because they were apart from each other at 13.4 angstrom, and as a result, the 3 SLH domain also moved its orientation (Figure 6.3.24. C). This amino acid change could be an important key for trypsin digestion of indigenous EA1.

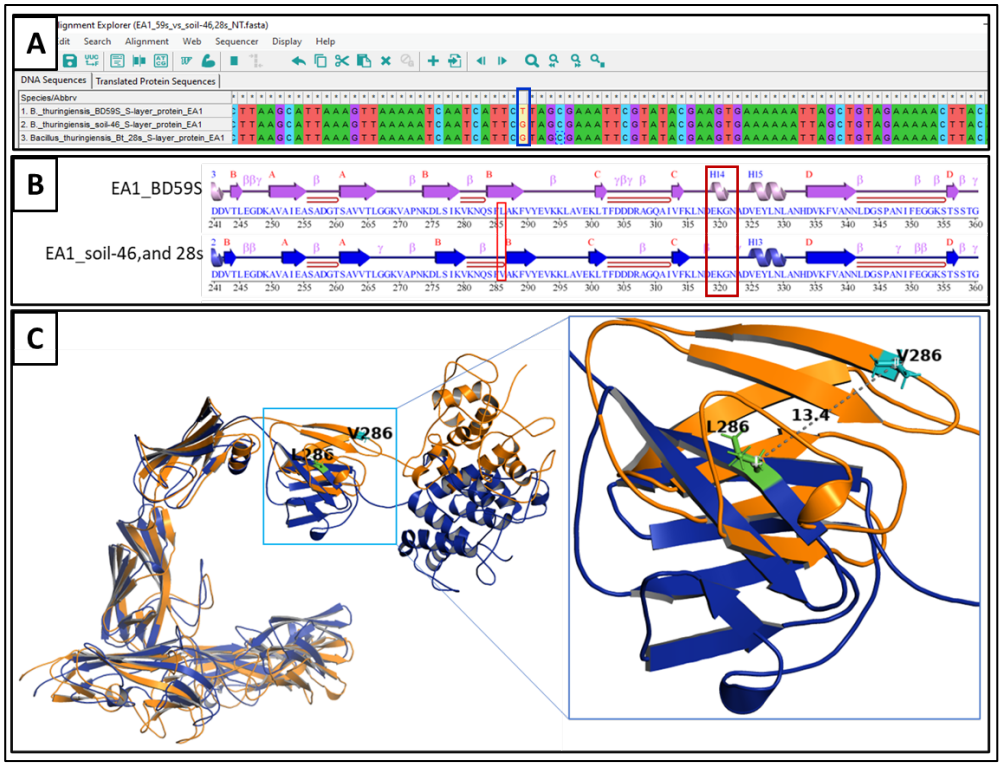



**Figure 6.3.24:** Structural difference of 91 kDa EA1 of BD59S and Soil-46 and 28S. (A) single nucleotide polymorphism; (B) comparative secondary protein structure with single amino acid change; (C) close view of comparative tertiary protein structure showing the structural difference in the protein of 91 kDa S-layer protein EA1 of indigenous Bt strains BD59S, Soil-46, and 28S.

#### 6.3.4.7.3 Domain study of the 73.54 kDa Cry-like protein of BD59S, Soil-46, and 28S

Besides, the S-layer protein, a 616 amino acid long 73.54 kDa hypothetical protein molecule, was thought to be synthesized as a parasporal protein in Bt strains BD59S, Soil-46, and 28S, because its size was thought to be the same as the SDS PAGE observed ~76 kDa protein band. We looked at the 73.54 kDa hypothetical protein using the InterPro scan, the NCBI protein features study, the NCBI domain search, and the simple modular architecture research tool (SMART). The 73.54 kDa hypothetical protein of the native Bt strains BD59S, Soil-46, and 28S did not have any well-identified domains. But as in the SDS PAGE analysis, nearly the same-sized parasporal protein band was found; hence, domain analysis was performed based on *in silico* tertiary and secondary protein structure prediction. From the Robetta predicted 3D protein tertiary structure, it was found that a three-domain Cry-like protein structure was formed from the 73.54 kDa protein sequences of indigenous Bt strains BD59S, Soil-46, and 28S (Figure 6.3.25. A), where domain-1 was a β-strand rich β-sandwich-like domain (from amino acids 1 to 163, red in the figure), domain-2 was α-helix rich domain (from 164 to 392, yellow in the figure), and domain-3 was a composition of both β-strand and α-helix (from amino acids 393-616, green in the figure). Besides, from the automatic fold recognition and protein structure prediction psi-blast results of the PHYRE2 web server, it was found that the 73.54 kDa protein had 40% structure identity (from amino acids 126–140) with a tumor necrosis factor signaling protein region. Besides, it also had a 21-27% structure identity with two different types of DNA-binding proteins (from amino acids 112–198 and 489–582). Besides, 35.1% structure identity (from 479 to 575) with a nucleoplasmin-like/VP (viral coat and capsid proteins) of Picornaviridae-like positive stranded ssRNA viruses and 20% structure identity (415–458) with the hica3 toxin of *Yersinia pestis.* Besides, from the pdbsum webserver predicted secondary structure, it was distinctly found that the N-terminal and the C-terminal portions of the protein molecule produced 2 β-strand rich domains, whereas the middle portion produced α-helices rich domains (Figure 6.3.25. C).

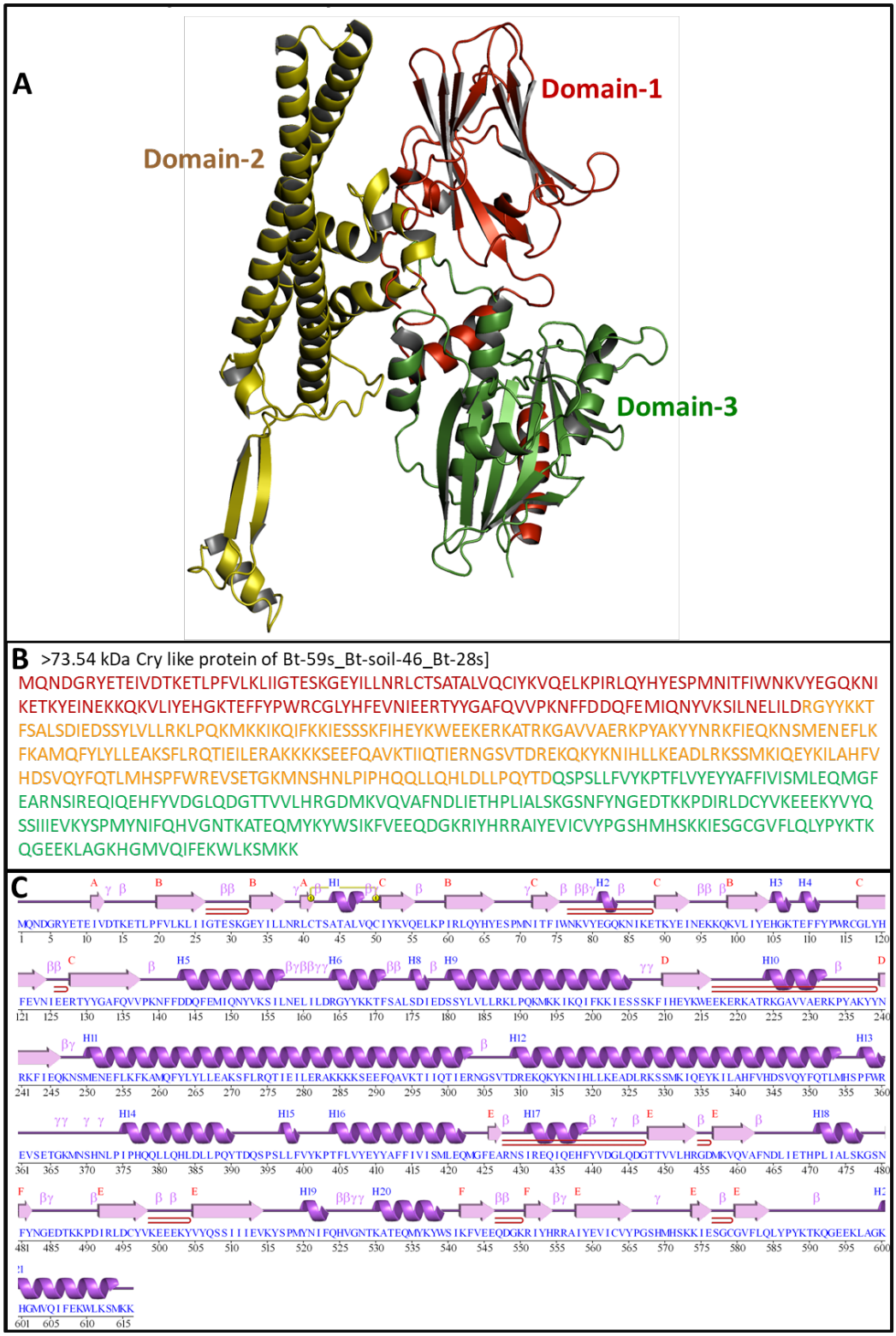



**Figure 6.3.25:** 3D and 2D protein structures of the 73.54 kDa Cry-like protein of indigenous Bt strains. (A) *In silico* 3D tertiary protein structure with predicted domains where domain 1 is red, domain 2 is yellow and domain 3 is green; (B) amino acid sequence colored to the domains; and (C) the secondary protein structure of 73.54 Cry-like protein.

### 6.3.4.7.4 Domain study of the 93 kDa S-layer protein of 45L

From the protein domain study, it was found that the 93 kDa S-layer protein of 45L had a signal peptide from amino acids 1-30 and 3 SLH domains, 33-78, 95-138, and 156-201 (Figure 6.3.26), and all of these SLH domains had α-helices and loops as well. Besides these, the 93 kDa SLP also had a 220 to 880 amino acid long β-strand-rich six β- sandwich crystallization or assembly domains (Figure 6.3.27. A). An aerolysin domain (472 to 779) was identified from the SMART search tool (Figure 6.3.26) and this domain was found within the β-strand-rich crystallization domain, and about 3 β-sandwich regions belonged to this domain (Figure 6.3.27. A). Besides, two carbohydrate-binding modules, CBM2 (575 to 662), and CBM3 (770 to 836) were also found within the long β-strand-rich crystallization domain of the 93 kDa S-layer protein (Figure 6.3.27. A), and of them, CBM2 had sequence similarities with aerolysin and hence lay within the aerolysin domain. Besides, two immunoglobulin-like copper resistance regions, CopC/internalin (219-320, 425-534), were also observed in the 93 kDa SLP from the InterPro domain scan. From the InterPro scan, it was found that the 93 kDa SLP was identical to Sap.

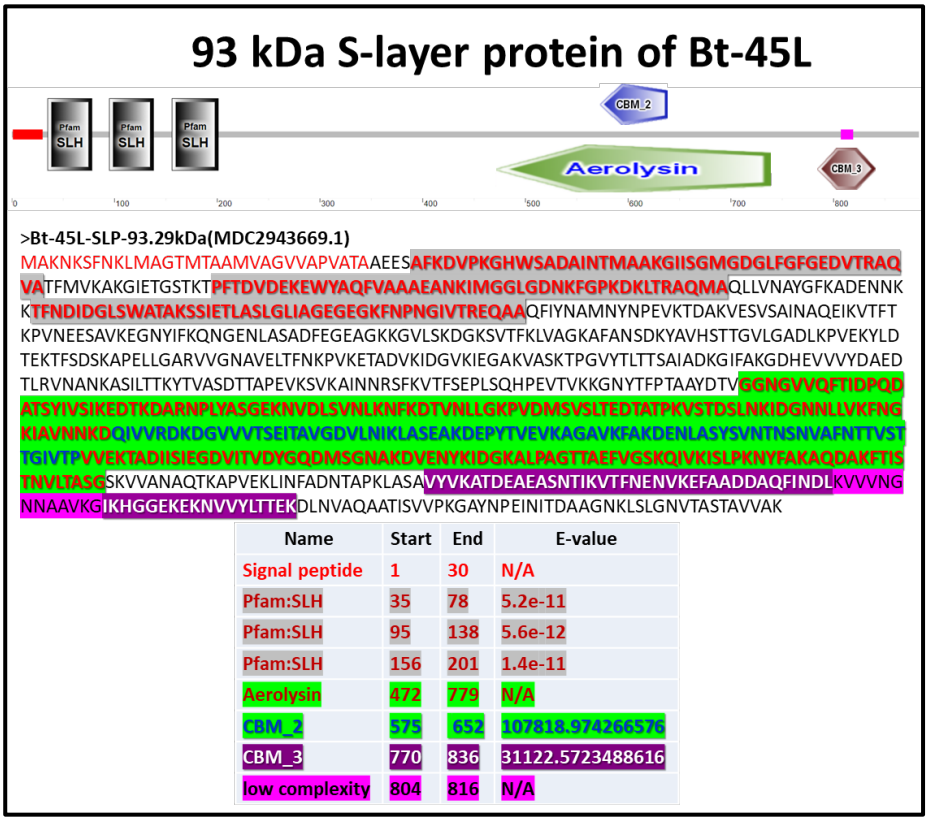



**Figure 6.3.26:** Schematic diagram of domains and features of the 93 kDa SLP of 45L.

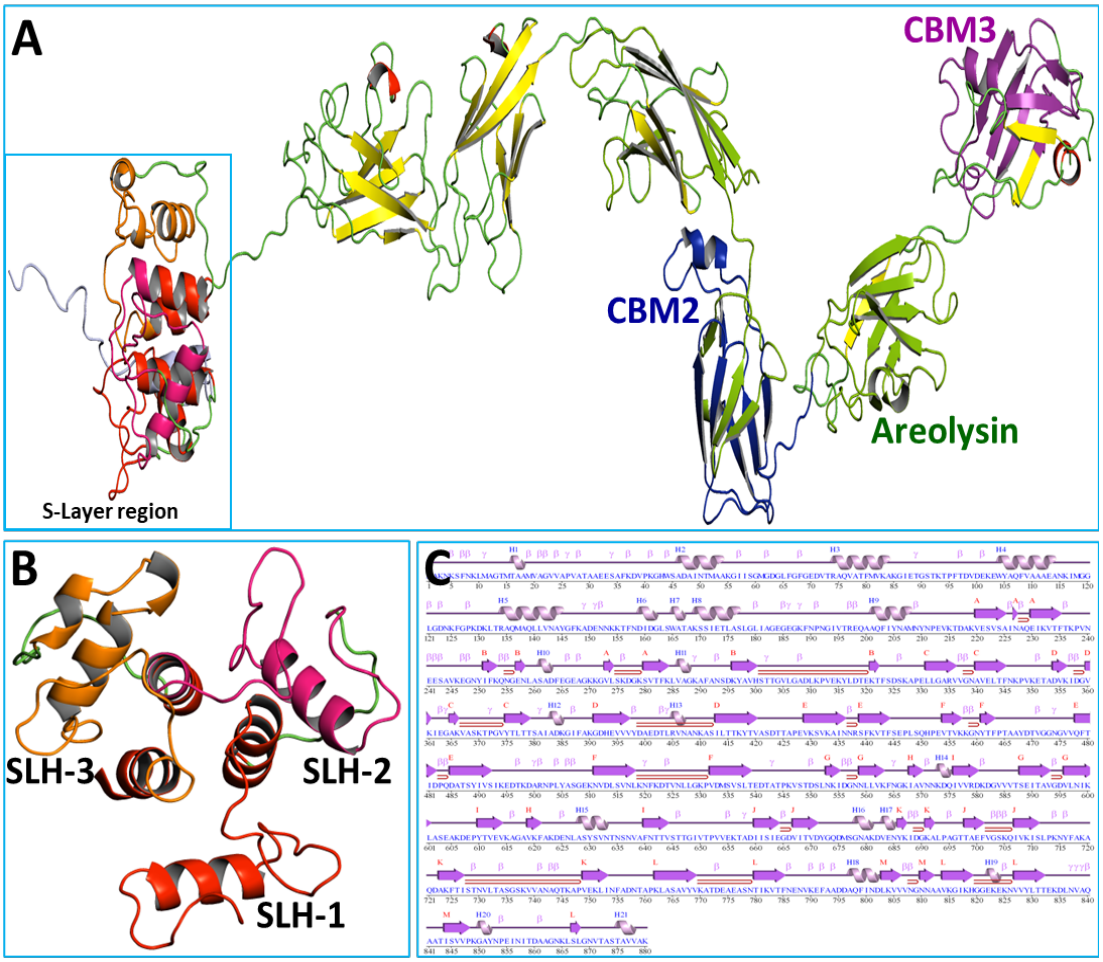



**Figure 6.3.27:** 3D and 2D protein structures of the 93 kDa S-layer protein of 45L. The tertiary protein structure in (A) shows SLH domains and an β-strand-rich crystallization domain. The secondary protein structure of a 93 kDa S-layer protein can be seen in (B) as a close-up of three SLH domains.

### 6.3.4.7.5 Domain study of the 103 kDa S-layer protein of 45L

The 103.11 kDa S-layer protein had a 27 amino acid long signal peptide and 6 domains, of them a MAEBL (merozoite apical erythrocyte-binding ligand) (28–309), 3 S-layer homology domains (299–340, 358–400, 420–462), a ricin-type beta trefoil (661–770), and a YvpB cysteine peptidase C39-like domain (708–901) (Figure 6.3.28; Figure 6.3.29 A). Besides, 2 β-strand-rich domains were also observed in the 3D tertiary and 2D secondary protein structures (Figure 6.3.29, A, C).

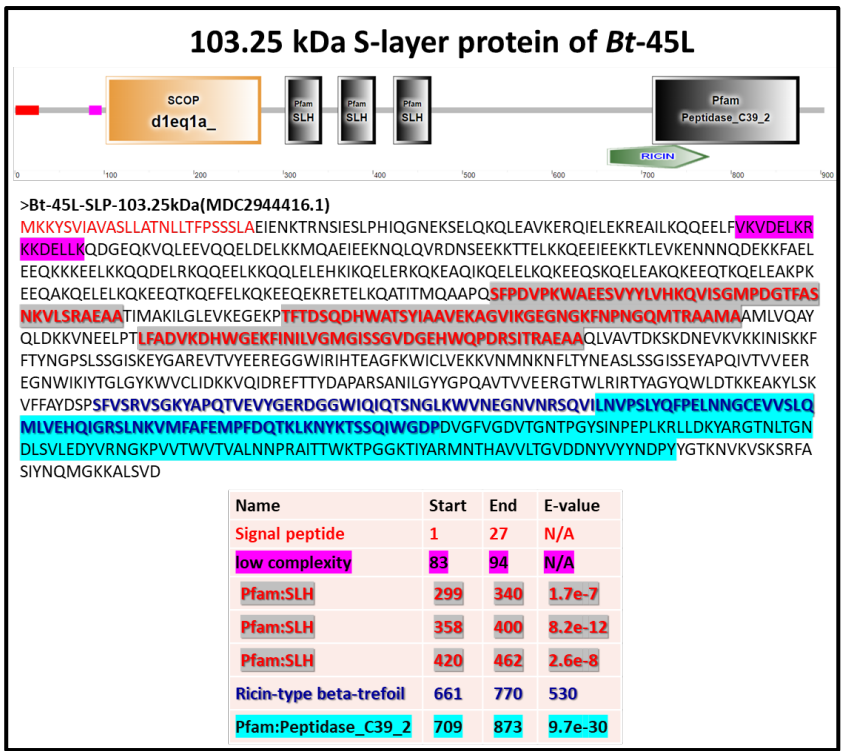



**Figure 6.3.28:** Schematic diagram of domains and features of the 103 kDa S-layer protein of 45L

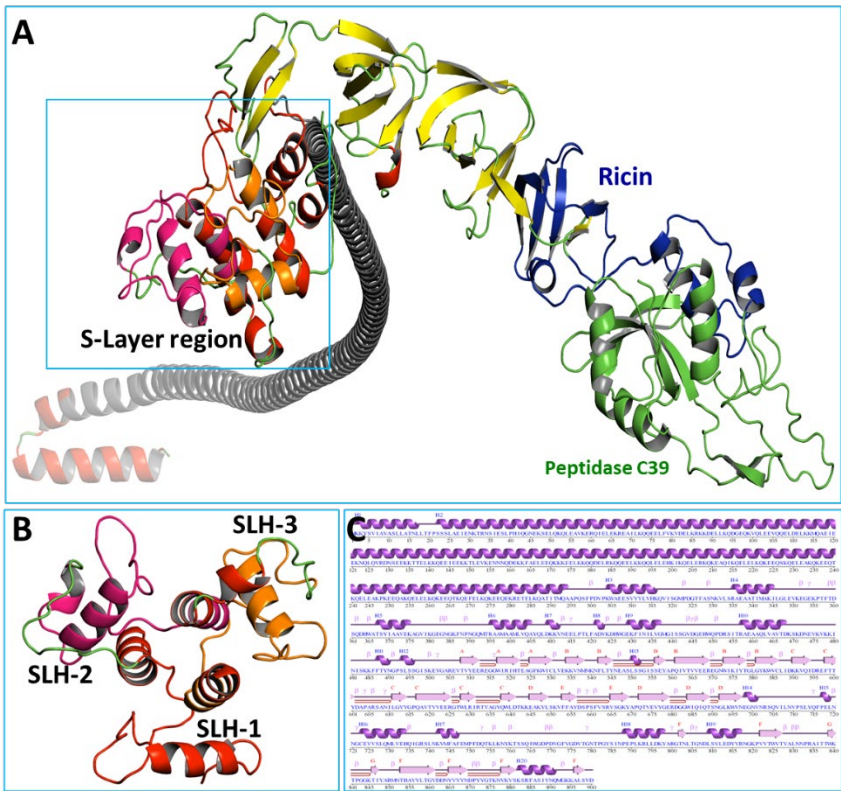



**Figure 6.3.29:** 3D and 2D protein structure of the103 kDa S-layer protein of 45L. (A) the Tertiary protein structure with SLH domains and β-strand-rich crystallization domains; (B) a close view of 3 SLH domains; and (C) the secondary protein structure of 103 kDa S-layer protein.

#### 6.3.4.8 Phylogenetic relationship and percentage of identity of indigenous SLPs with anticancer parasporins and insecticidal Cry

As the anticancer activity was found from the proteinase K digested and trypsin digested parasporal proteins of 5 indigenous Bt strains and 4 different sizes parasporal S-layer protein (91 kDa, 86 kDa, 103 kDa, and 93 kDa) and 73.54 kDa Cry like proteins were identified from the genome sequence of the indigenous Bt strains, therefore phylogenetic tree, as well as the percentage of identity of these indigenous proteins between anticancer parasporins and insecticidal Cry toxins, were conducted.

The amino acid sequences of indigenous probable parasporal SLP, Cry-like 73.54 kDa protein, and six different classes of parasporins were used to make a phylogenetic tree. From the phylogenetic tree of indigenous SLPs, cry-like proteins, and six different classes of parasporins (Figure 6.3.30), it was found that six different types of parasporins were divided into two major nodes, and indigenous SLPs and the 73.54 kDa cry-like proteins were also placed in two separate nodes. It was distinctly found that the upper positioning node was once again divided into two distinct and separate clades, where all indigenous SLPs created a separate clade, and PS2, PS4, and PS5 altogether created another separate clade. The clade contained PS2, PS4, and PS5, again divided into two separate taxa, where PS2 produced a single closely related taxon and PS4 and PS5 produced another close taxon. From this node and clades, it was clear that indigenous SLPs have some sequence similarities and relationships with the necrotic pore-forming type parasporins (PS2, PS4, and PS4). On the other hand, another node was also divided into two major clades, where indigenous Cry-like 73.54 kDa protein created a single clade as well as a single taxon. Besides, PS1, PS3, and PS6 again created two distinct clades, where PS6 created a single clade as well as a single taxon, and PS1 and PS3 produced two different taxa. From this node and clades, it was clear that the indigenous Cry-like 73.54 kDa protein has some sequence similarities and relationships with the apoptotic cancer cell-killing type parasporins (PS1, PS3, and PS5).

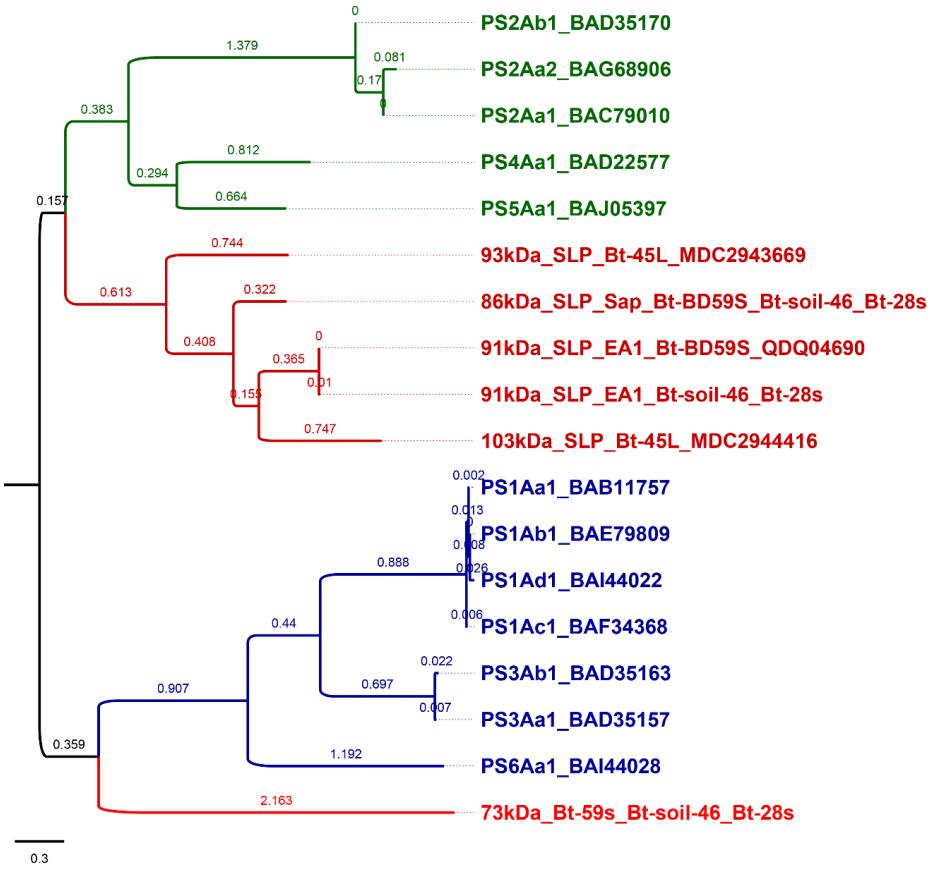



**Figure 6.3.30:** Phylogenetic relationship of indigenous SLPs and Cry-like 73.54 kDa protein with parasporins. The accession number is at the end of each taxon name.

The amino acid sequences of native probable parasporal SLP, Cry-like 73.54 kDa protein, and sorted Cry proteins were also used to make a phylogenetic tree. At the beginning of this analysis, indigenous SLPs and 73.54 kDa proteins were aligned all the Cry protein sequence found in the Cry protein database (old website: http://www.lifesci.susx.ac.uk/home/Neil_Crickmore/Bt/; new website: https://www.bpprc-db.org/home/), and then a phylogenetic tree was constructed using the alignment data. From that phylogenetic tree, closely related Cry proteins were selected and this final version of the phylogenetic tree was constructed. From the phylogenetic tree of indigenous SLPs, the 73.54 kDa cry-like protein and sorted cry (Figure 6.3.31), it was found that two nodes were produced. Among these 2 nodes, SLPs were located in node 1 (upper portion of the tree) and the 73.54 kDa Cry-like protein was located in node 2 (lower portion of the tree). Each of the nodes was divided into two major clades. It was found that the upper positioning clade once again divided into two distinct and separate sub-clades, where all indigenous SLPs created a separate subclade, and cry0034 (accession: AEB52307, a parasporin-like protein) and a delta-endotoxin (accession: ACN87260) altogether created another separate subclade. On the other hand, the 73.54

kDa cry-like protein-containing clade of node 2 is again divided into subclades. It was found from the tree that the 73.54 kDa cry-like protein had significant sequence similarities with a cry toxin of *Lysinibacillus sphaericus* (accession: CAH56541) and a Bt cry toxin (accession: AAK64558). Besides, the 73.54 kDa protein also had similarities with the two-domain toxin of Bt israelensis (accession: CAD30104) and a Cyt-like Bt toxin (accession: HQ113115) located in the other subclade of it.

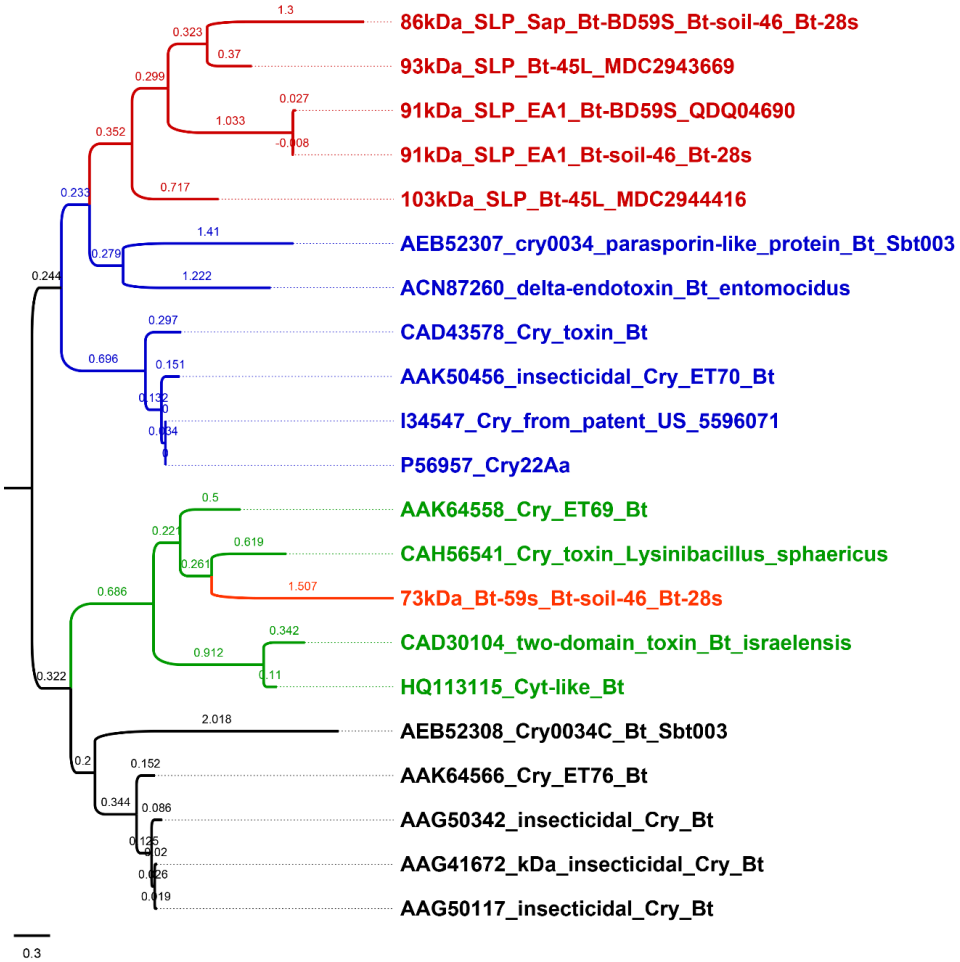



**Figure 6.3.31:** Phylogenetic relationship of indigenous SLPs and Cry-like 73.54 kDa protein with Cry toxin. The accession number is at the beginning of each taxon name.

From the percentage of identity study, it was found that the 91 kDa EA1 and 86 kDa Sap of indigenous strains BD59S, 28S, and Soil 46 were non-Cry family S-layer proteins, and they had 19% sequence similarities with anticancer Cry protein parasporin-3, PS3Ab1 (BAD35163.1) and PS3Aa1 (BAD35157.1) (Figure 6.3.32. A-B). Besides, the Cry-like 3 domain 73.54 kDa hypothetical protein of the 3 indigenous strains also had 18% sequence homologies with cancer cell-killing Cry protein parasporin-3, PS3Ab1 (BAD35163.1) and PS3Aa1 (BAD35157.1). In addition, the 73.54 kDa protein also had 15.38% and 14.59% sequence similarities with two

insecticidal proteins, a delta-endotoxin of Bt entomocidus (ACN87260) and a Cry toxin of *Lysinibacillus sphaericusa* (CAH56541), respectively (Figure 6.3.32. C). Besides, the 93 kDa SLP of 45L had 16% and 15% amino acid sequence similarities with anticancer PS3Ab1 (BAD35163.1) and PS3Aa1 (BAD35157.1), respectively. Moreover, the 93 kDa SLP also had 20.89% sequence similarities with the crystal toxin of *L. sphaericus* (CAH56541.1), 23.29% with the insecticidal crystal protein CryET70 of Bt (AAK50456.1), 22.30% with the pesticidal crystal protein Cry22Aa (P56957.1) of Bt, and 15.72% with cry0034, a parasporin-like protein of Bt strain Sbt003 (AEB52307) (Figure 6.3.32. D). It was also found that the amino acid sequence of the 103 kDa SLP of the indigenous 45L strain had 14.22% amino acid sequence similarities with cancer cell-killing Cry protein PS5Aa1 (BAJ05397), 19.84% with PS3Aa1 (BAD35157.1), and 19.35% with PS3Ab1 (BAD35163.1) (Figure 6.3.32. E).

**A**

# Percent Identity Matrix - created by Clustal2.1

| | | | | |
|---|---|---|---|---|
| 1: 91kDa_SLP_EA1_Bt-BD59S_QDQ04690] | 100.00 | 99.88 | 19.04 | 19.54 |
| 2: 91kDa_SLP_EA1_Bt-soil-46_Bt-28s] | 99.88 | 100.00 | 19.04 | 19.54 |
| 3: PS3Aa1_BAD35157 | 19.04 | 19.04 | 100.00 | 89.79 |
| 4: PS3Ab1_BAD35163 | 19.54 | 19.54 | 89.79 | 100.00 |

**B**

# Percent Identity Matrix - created by Clustal2.1

| | | | |
|---|---|---|---|
| 1: 86kDa_SLP_Sap_Bt-BD59S_Bt-soil-46_Bt-28s] | 100.00 | 19.94 | 19.97 |
| 2: PS3Aa1_BAD35157 | 19.94 | 100.00 | 89.79 |
| 3: PS3Ab1_BAD35163 | 19.97 | 89.79 | 100.00 |

**C**

# Percent Identity Matrix - created by Clustal2.1

| | | | | | |
|---|---|---|---|---|---|
| 1: 73kDa_Cry_like_Bt-59s_Bt-soil-46_Bt-28s] | 100.00 | 18.02 | 18.48 | 15.38 | 14.59 |
| 2: PS3Ab1_BAD35163 | 18.02 | 100.00 | 89.79 | 19.76 | 11.43 |
| 3: PS3Aa1_BAD35157 | 18.48 | 89.79 | 100.00 | 19.88 | 11.14 |
| 4: ACN87260_delta-endotoxin_Bt_entomocidus] | 15.38 | 19.76 | 19.88 | 100.00 | 21.59 |
| 5: CAH56541_Cry_toxin_Lysinibacillus_sphaericus] | 14.59 | 11.43 | 11.14 | 21.59 | 100.00 |

**D**

# Percent Identity Matrix - created by Clustal2.1

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1: 93kDa_SLP_Bt-45L_MDC2943669 | 100.00 | 16.06 | 15.09 | 20.89 | 23.29 | 22.30 | 15.72 |
| 2: PS3Ab1_BAD35163 | 16.06 | 100.00 | 89.79 | 10.63 | 14.40 | 15.42 | 25.73 |
| 3: PS3Aa1_BAD35157 | 15.09 | 89.79 | 100.00 | 11.92 | 14.34 | 14.95 | 26.16 |
| 4: CAH56541_Cry_toxin_Lysinibacillus_sphaericus | 20.89 | 10.63 | 11.92 | 100.00 | 12.05 | 12.95 | 14.49 |
| 5: AAK50456_insecticidal_Cry_ET70_Bt | 23.29 | 14.40 | 14.34 | 12.05 | 100.00 | 88.35 | 16.36 |
| 6: P56957_Cry22Aa] | 22.30 | 15.42 | 14.95 | 12.95 | 88.35 | 100.00 | 15.80 |
| 7: AEB52307_cry0034_parasporin-like_protein_Bt | 15.72 | 25.73 | 26.16 | 14.49 | 16.36 | 15.80 | 100.00 |

**E**

# Percent Identity Matrix - created by Clustal2.1

| | | | | |
|---|---|---|---|---|
| 1: 103kDa_SLP_Bt-45L_MDC2944416] | 100.00 | 14.22 | 19.84 | 19.35 |
| 2: PS5Aa1_BAJ05397 | 14.22 | 100.00 | 19.68 | 20.16 |
| 3: PS3Aa1_BAD35157 | 19.84 | 19.68 | 100.00 | 89.79 |
| 4: PS3Ab1_BAD35163 | 19.35 | 20.16 | 89.79 | 100.00 |

**Figure 6.3.32:** Percent Identity Matrix of S-layer and Cry-like 3-domain proteins of indigenous Bt strains with anticancer parasporins and insecticidal Cry toxins. (A) 91 kDa EA1; (B) 86 kDa Sap; (C) 73.54 kDa Cry-like protein; (D) 93 kDa S-layer protein; and (E) 103 kDa S-layer protein. (Created by Clustal2.1 of EBI Clustal Omega multiple sequence alignment)

# CHAPTER 7:
# DISCUSSION AND CONCLUSION

# CHAPTER 7: DISCUSSION AND CONCLUSION

## 7.1 Discussion

Parasporal Cry proteins that are nonhemolytic, and noncytotoxic to normal cell lines but are capable of killing human cancer cells are called parasporins. Japanese scientist Mizuki and his team made a groundbreaking discovery in Bt research by identifying three non-insecticidal parasporal Cry proteins from nonhemolytic Bt strains that exhibited targeted cytotoxicity against human leukemia cells (Mizuki *et al.*, 2000). The discovery prompted a worldwide search for non-insecticidal strains of *B. thuringiensis* in order to locate parasporal anticancer Cry proteins.

In addition to parasporal Cry protein, *B. thuringiensis* also produced another type of parasporal inclusion protein, called surface layer protein (SLP). In 2001, for the first time, a 100 kDa surface layer protein (SLP) was identified to produce atypical oval-shaped proteinaceous parasporal inclusion in the *B. thuringiensis* strain CTC (Sun *et al.*, 2001). Afterwards, it was demonstrated that *B. thuringiensis* strains CTC, B22, I13, H67, and BMB1152 synthesized two distinct forms of parasporal SLPs, namely the surface array protein (Sap) and the extractable antigen-1 (EA1) (Zhou *et al.*, 2011). Like parasporins, the cytotoxic contribution of SLP of Bt to cancer cells was recently discovered. The 86 kDa SLP of *B. thuringiensis* strain AP11 was found to be cytotoxic on MDA-MB-231 Breast Cancer Cells, but noncytotoxic to HaCat non-cancer cell lines (Rubio *et al.*, 2017). In addition to Bt, recent research showed that SLPs of lactic acid bacteria (LAB) decrease malignant tumour development and spread and prevent colon cancer (Wu *et al.*, 2019). The SLP of LAB could induce apoptosis in HT-29 cells via both the death receptor apoptotic pathway and the mitochondrial pathway (Figure 2.35), (Zhang *et al.*, 2020). The SLP derived from *L. acidophilus* CICC 6074 exhibited cytotoxic effects on colon cancer HT-29 cells (Liu *et al.*, 2021). So, parasporins as well as SLP can be anticipated to leave a mark in cancer research.

### 7.1.1 Isolation and characterization of novel genes from *B. thuringiensis* strains: Implications for cancer cell targeting

Human cancer cell killing parasporin-1 (Cry31A), and parasporin-6 (Cry63Aa) were isolated from nonhemolytic *B. thuringiensis* strains (Ichikawa *et al.*, 2008; Nagamatsu *et al.*, 2010), and parasporin-2 (Cry46Aa), parasporin-3 (Cry41Aa and Cry41Ab), and

parasporin-4 (Cry45Aa) were isolated from hemolytic Bt strains (Ichikawa *et al.*, 2008). So, it is apparent that both hemolytic and nonhemolytic Bt strains have the potential to possess the parasporin genes. Considering this, an investigation was conducted to assess the hemolytic capabilities of the previously identified two hundred ninety-one (291) indigenous *B. thuringiensis* strains. As a result, a total of 98 (34%) were identified as nonhemolytic whereas 193 (66%) identified as hemolytic Bt strain (Figure 3.3.2).

All 98 non-hemolytic (Table 3.2.1) strains and 23 hemolytic strains (Table 3.2.2) were then selected to investigate the presence of parasporin genes (*ps1-ps5*) through PCR using the primer pairs, constructed newly from the conserved sequence of the existing parasporin database. Most of the *cry* genes are located in the plasmid DNA of Bt stains (Gonzalez *et al.*, 1980), hence, both plasmid and genomic DNA were examined in this experiment. The obtained PCR amplicons were then submitted to DNA sequencing and subsequent data analysis were performed.

Parasporins are very rare in environment, and till to date only 19 parasporins are identified and enlisted in the parasporin database (https://www.fitc.pref.fukuoka.jp/parasporin/list.html). In our experiment, a total of 27 PCR amplicons were obtained from 121 indigenous strains, of them 20 from nonhemolytic (Table 3.3.1) and 7 from hemolytic strains (Table 3.3.2). Among 27 PCR amplicons, DNA sequencing was performed for 25 PCR products. There were 25 contigs generated from the 25-sequencing data. This indicated that complete PCR had took placed from each set of primers. So, to know the gene name, as well as protein name and sequence, the NCBI BLAST was performed. Regrettably, the BLAST findings did not align with any of the known parasporins or cry genes documented in the database. A concise overview of the identified proteins including their molecular weight, source, and accession numbers were presented in Table 3.3.3 and Table 3.3.4.

The 550 bp PCR amplicon of *ps1* specific primers from nonhemolytic strain FhSb3, were considered to be parasporin-1 positive, but from DNA sequencing data it was identified as group II intron reverse transcriptase/maturase enzyme synthesizing gene. Besides, the *ps3* (*CryBP*) gene-specific PCR primers successfully generated 7 PCR amplicon of the anticipated size (499 bp) from plasmid DNA of both nonhemolytic and hemolytic strains, but from DNA sequencing data, a protein-glutamate O-methyltransferase (Chemotaxis protein) synthesizing gene *CheR* was identified from

nonhemolytic strains, when a polysaccharide deacetylase family protein synthesizing gene was identified from hemolytic strains. In this case, although the size of PCR amplicons from both hemolytic and nonhemolytic strains were identical, but their identity was different. Besides, the primers for *ps3* also amplified a 600 bp amplicon from reference Bt strain *sotto* and that was also a multidrug/ EamA family transporter. The primers for *ps4* (*etx-mtx*) amplified a Tn3 family transposase from five nonhemolytic Bt strains (Figure 3.3.5 A, B), whereas same primers amplified a gene that encode DMT family transporter from two hemolytic strains. There is one notable point, the *ps4* and *ps3* specific primers amplified two different transporter genes from three hemolytic strains (Table 3.3.4), and *ps1* primer also amplified a transporter gene from two nonhemolytic strains (Table 3.3.3).

The primers for *ps4* (*etx-mtx*) amplified a Tn3 family transposase from five nonhemolytic Bt strains. The Bt strains possess a wide range of transposable elements, such as insertion sequences and transposons. *B. thuringiensis subsp. israelensis* contains six endotoxins that are found in the parasporal crystals. Among these genes, at least two are connected to transposable elements: cry4A, which is bordered by two copies of IS240 (IS240A and IS240B), and IS231W, which is positioned next to cry4D (Mahillon *et al.*, 1994). The cry11Aa gene in Bt subsp. israelensis is located next to a transposable element called IS231W, and the IS231 elements are commonly linked to cry genes. So, the amplified transposable elements in these experiments, also could have some relations with Cry like gene.

The DMT and EamA superfamily transporters are responsible for exporting a diverse array of substances, including toxic substances and metabolites. In addition, EamA is the sole drug/metabolite transporter family that has the ability to traverse the boundary between prokaryotes and eukaryotes (Jack *et al.*, 2001). It is known that PS2, PS3, PS4, and PS5 kills cancer cells by pore formation hence these PSs have pore forming domains. As three different primers amplify DMT/EamA family transporter genes from both hemolytic and nonhemolytic. DMT/EamA like transporters also penetrates the membrane boundary, so, there could be a relationship between these types of transporters with pore forming PSs.

From this experiment, unfortunately, no match was observed between the sequences and the known parasporin genes in the existing database, but we identified 11 of distinct

types of genes on the plasmids. The genes amplified by parasporin-specific primers primarily consist of genes that play a crucial role in controlling cell growth and maintenance. These include genes encoding bacterial membrane-related transporter proteins, bacterial transposable elements, and enzymes involved in the manipulation of DNA and RNA. Therefore, it is hypothesized that these genes from *B. thuringiensis* may also have a role in killing human cancer cells, like parasporins.

### 7.1.2 Investigating parasporal protein diversity in nonhemolytic *B. thuringiensis* strains (through SDS-PAGE analysis)

Failure to obtain definite parasporin genes through genomic study led us to effectively implement proteomics. Hence, extraction and purification of parasporal proteins from 98 nonhaemolytic strains and their SDS-PAGE analysis were performed. Seven strains namely MyIa2, FHSc4, KkSc1, RaSd1, MyIb1, MyIa1, and SgSp2 were found to produce one or more parasporal proteins with diverse molecular mass ranging from 17 kDa to 78 kDa.

Parasporins can be categorized into two groups based on their molecular mass. The first group includes PS1, PS3, and PS6, which have a higher molecular mass of around 80 kDa. The second group consists of PS2, PS4, and PS5, which have a lower molecular mass ranging from 30 to 37 kDa. In this study, molecular masses of some parasporal proteins on SDS-PAGE somewhat corresponded to the known parasporins, PS1, PS2, PS4, and PS5. A 78 kDa parasporal protein was found from the SgSp2, which was close to the 81 kDa PS1 (Mizuki *et al.*, 2000; Uemori *et al.*, 2008; Yasutake *et al.*, 2008; Kuroda *et al.*, 2013). Besides, a 28 kDa protein band from MyIa1, MyIb1, and a 30 kDa band from MyIa1, RaSd1, were close to the 30kDa PS2Aa2 (Okumura *et al.,* 2013). Besides, 32 kDa protein bands were found from FHSc4 and MyIa2, which were close to the 30 kDa PS4Aa1 (Okumura *et al.*, 2004), and the 33 kDa PS2Ab1, and PS5Aa1. These findings indicated the possible presence of PS1, PS2, PS4, and PS5 in indigenous nonhemolytic Bt strains. Moreover, protein bands around 56 kDa to 58kDa were visible from RaSd1, MyIb1, FHSc4, and KkSc1 strains; and protein bands around 40 kDa to 46 kDa were visualized from MyIa1, MyIa2, FHSc4 and KkSc1 strains. These, along with a few other unexpected-sized proteins, were under investigation with the aim to find novel parasporins.

An interesting feature of parasporins are that, like other Cry proteins, they have to be activated through a proteolytic cleavage by either proteinase K or Trypsin (Knowles, 1994; Brasseur *et al.,* 2015). In this perspective, parasporal protein of each of seven strains were digested with proteinase K and then proteinase K activated parasporal proteins were tested to detect their cytotoxicity on HeLa cell lines. As a result, among seven nonhemolytic Bt strains, proteinase K-treated solubilized parasporal proteins of three strains, MyIa2, FHSc4, and KkSc1 were found to be cytotoxic to the HeLa cell line. The 78kDa parasporal protein of SgSp2 that had a molecular mass close to the 81 kDa parasporin (PS1Ac2), as well as parasporal proteins from the other three strains RaSd1, MyIb1, MyIa1 didn't show any cytotoxic effect on the HeLa cell line.

Although, proteinase K digested parasporal protein of three strains, MyIa2, FHSc4, and KkSc1 showed cytotoxicity to the HeLa cell line, but unexpectedly, none of the putative parasporal proteins obtained from these three strains along with other four strain SgSp2, RaSd1, MyIb1, and MyIa1 produced any visible protein fragments on SDS-PAGE upon proteinase K treatment. This finding contradicts the fact that our observed proteins were parasporin belonging to parasporin families PS1-PS6.

Then, to confirm the source of SDS-PAGE observed parasporal protein bands were really from the solubilized form of parasporal inclusion of the strains, n-hexane purified Crystals of MyIa2, FHSc4, KkSc1, along with the reference Bt strain HD-73 were observed under phase-contrast microscope. Purified parasporal inclusion bodies under phase-contrast microscope were revealed as round in KkSc1, cubic and bipyramidal shape in reference strain HD73, and indigenous MyIa2 and FHSc4. Microscopic observation of parasporal inclusions supported that the previously observed protein bands from SDS-PAGE of indigenous Bt strains were the solubilized form of parasporal protein. *B. thuringiensis* produces several uncharacterized parasporal inclusions. A few of them may possess anticancer, or insecticidal or poisonous properties towards animals that have not undergone testing. The fact that a *B. thuringiensis* strain utilizes significant resources and energy for the formation of these inclusions, implies that they have some function.

The main goal of this research was to find the parasporin protein that can suppress cancer cells from native nonhemolytic *B. thuringiensis*. Through this study, we have identified parasporal proteins-producing indigenous nonhemolytic *B. thuringiensis*

strains. Although some of the proteins we identified have similar sizes to PS1, PS2, PS4, and PS5, none of these parasporal proteins can be called parasporins or any other cry proteins because they are completely digested by proteinase K. As the proteins being studied do not belong to the parasporins or cry proteins, their ability to cause cell death in human cancer cells prompts further investigation into their unidentified biological nature.

### 7.1.3 Uncovering novel parasporal proteins with potential anticancer properties in indigenous hemolytic *B. thuringiensis* strains

Parasporal inclusions are mainly different types of Cry, Cyt, S-layer, and unknown proteins, and each has a different identifiable shape. The proteinase K or trypsin-digested solubilized parasporal protein of *B. thuringiensis* exhibits a specific cytotoxic effect on cancer cells, which could be significant in developing effective protein-based anticancer treatments. Therefore, a promising opportunity arises to develop protein-based therapies for cancer treatment by utilizing the anticancer properties of parasporal inclusion proteins derived from Bt. Reports are also found on anticancer Cry proteins PS2, PS3, and PS4 isolated from hemolytic Bt strains (Ichikawa *et al.*, 2008). This study was thus conducted to discover the novel parasporal anticancer proteins from previously identified twenty-three indigenous non-insecticidal hemolytic Bt strains.

Five indigenous Bt strains, BD59S, Soil-46, 28S, 45L, and Dsh4 were found to produce parasporal protein band on the SDS-PAGE (Figure 5.3.2), of them, BD59S, Soil-46, 28S were found to produce 3 proteins of identical molecular mass (101 kDa, 86 kDa, and 76 kDa) (Figure 5.3.2. A). Besides, Dsh4 and 45L were found to produce a single and identical ~103 kDa protein band in the SDS-PAGE (Figure 5.3.2. B). The 103 kDa (from Dsh4 and 45L) and 76 kDa (from BD59S, Soil-46, 28S) was in agreement with the previously reported HeLa cells, human leukemic T cells (MOLT-4) and the normal T cells killing *B. thuringiensis* strains 90-F-45-1 (new name A1547) having identical or somewhat close molecular mass at least two of five (170, 103, 73, 40, and 32 kDa) parasporal inclusion proteins (Kim *et al.,* 2000). Besides, the 86 kDa parasporal inclusion protein of BD59S, Soil-46, and 28S strains exhibited very close similarity in size with the parasporin PS1Ac1(86.76kDa), PS3Aa1(88kDa), PS3Ab1(88kDa), and

PS6Aa1(85 kDa) from *B. thuringiensis* strains Bt 87-29, A1462, CP84, M019, respectively (Yasutake *et al.,* 2008; Yamashita *et al.,* 2005; Nagamatsu *et al.,* 2010). Besides, parasporins, the presence of ~101 kDa and ~86 kDa parasporal protein bands of indigenous Bt strains BD59S, 28S, and Soil-46 were also in agreement with the size from the three different Bt strains GP1, CTC, and BMB1152 which are parasporal S-layer proteins (SLP) (Pena *et al.,* 2005, Lormendez *et.al.,* 2019, Zhou *et.al.,* 2011). An 86 kDa surface layer protein isolated from Bt strain AP11 which was known to have cytotoxic activity against MDA-MB-231 Breast Cancer Cells and noncytotoxic to HaCat non-cancer cell lines (Rubio *et.al.,* 2017). Hence, the 103 kDa, 101 kDa, and 86 kDa parasporal proteins of indigenous Bt strains were studied carefully to determine whether they were from PS or SLP. As our main focus was to identify the parasporin hence, we proceeded our journey focusing on the parasporal anticancer Cry protein.

Solubilized parasporal proteins from each of the five *B. thuringiensis* strains were digested with trypsin to be activated for cancer cell line and observed from SDS-PAGE. As a result, 64, 51, 45, 41 and 32 kDa (Figure 5.3.3) trypsin digested fragments were found from 28S and Soil 46-S, when 64, 51 and 41 kDa from strain BD59S. From tryptic digestion, 45 and 32 kDa fragments were not produced in BD59S. It means, any one of the 101, 86, or 76 kDa protein of BD59S have uncommon tryptic digestion site. Besides, a 45 kDa single fragment was found from Dsh4 and 45L (Figure 5.3.3). Tryptic digested 45 kDa fragment was also found in 28S and Soil-46. It means, the type or nature of the 103 kDa protein is identical to any one of the 101, 86, or 76 kDa protein.

In this experiment, the size of trypsin-digested protein fragments did not correspond to the size of any trypsin-digested fragments of the known parasporins. The records of activation of parasporin upon trypsin digestion was found only from PS1 and PS6. Trypsin-digested PS1 is known to produce 55 to 56 kDa protein fragments, while trypsin-digested PS6 is known to produce a 59 kDa protein fragment. Thus, it was believed that none of these parasporal proteins from native Bt strains belonged to the PS1 or PS6 class.

Like trypsin digestion, solubilized parasporal proteins from each of the five *B. thuringiensis* strains were also digested with proteinase K to become activated for cancer cell line and also to observe on SDS-PAGE. Following digestion, a single protein fragment with a molecular weight of about 46 kDa was detected in the SDS-PAGE

analysis for each of the five strains (Figure 5.3.4). This suggests that the 103 kDa parasporal protein of Dsh4 and 45L also have sequence identity with at least one of the 101, 86, or 76 kDa parasporal proteins of BD59S along with 28S and Soil-46.

In this study, our obtained proteinase K digested protein fragments didn't match with any one of proteinase K digested fragments of known parasporins. The reported Proteinase K digested parasporins are of different sizes. PS1Ac2 protein fragments are 15 and 60 kDa (Kuroda *et al.*, 2013), while PS2 produced fragments ranging from 28 to 30 kDa after the same treatment (Ito *et al.,* 2004; Okumura *et al.*, 2013; Hayakawa *et al.*, 2007). Besides, PS3 is known to produce a single fragment, 64 kDa in size (Yamashita *et al.,* 2005). PS4 and PS5 are reported to produce fragments of 28 kDa and 30 kDa, respectively (Okumura *et al.*, 2004; Ekino *et al.*, 2014). Therefore, it was hypothesized that none of these parasporal proteins derived from 5 indigenous Bt strains would be classified in the same classes with the known parasporins PS1 to PS5.

The solubilized and untreated crude parasporal inclusion proteins from five native Bt strains had *in vivo* cytotoxicity on the brine shrimp (*Artemia* nauplii), with an LC50 value of 114.192 µg/mL. Besides, the solubilized and untreated crude parasporal protein samples all five indigenous Bt strains were found to be nonhemolytic on sheep red blood cells and non-insecticidal to the Tephritidae fruit fly larvae (*B. cucurbitae*). The obtained non-hemolytic, non-insecticidal cytotoxic proteins resembled the characteristics of parasporal proteins.

Three distinct variants of solubilized parasporal inclusion protein samples, namely untreated, proteinase K treated, and trypsin treated, were individually examined from each of the 5 Bt strains to assess their cytotoxic effects on HeLa and Vero cell lines. Out of the 5 Bt strains, the untreated solubilized parasporal proteins of the BD59S did not show any cytotoxic effects on the HeLa cell line (Figure 5.3.8), whereas 28S, Soil-46, 45L, and Dsh4 showed weak cytotoxicity. Previous research reports have indicated that untreated parasporins did not exhibit cytotoxic effects on the HeLa cancer cell line (Mizuki *et al.,* 2000; Katayama *et al.,* 2006). However, untreated surface layer proteins demonstrated mild cytotoxicity towards the MDA-MB-231 breast cancer cell line (Rubio *et al.,* 2017) and the HT-29 colon cancer cell line (Zhang *et al.,* 2020, 2021).

The trypsin-digested parasporal inclusion proteins of the BD59S exhibited significant cytotoxicity towards the HeLa cancer cell line (Figure 5.3.9), whereas 28S, Soil-46, 45L, and Dsh4 showed only moderate cytotoxicity towards the HeLa cancer cell lines. BD59S exhibited more cytotoxicity than other strains, even though the crude protein of BD59S at the same concentration did not display any cytotoxic effects. The presence of enlarged and spherical deceased HeLa carcinoma cells suggests that they expired due to necrosis. It was reported that, trypsin-digested PS6 caused cell death in human hepatocyte cancer HepG2 cells and uterus cervix cancer HeLa cells by inducing cell swelling or the formation of vacuoles in the cytoplasm (Nagamatsu *et al.,* 2010), whereas trypsin-digested PS1 exhibited cytotoxicity towards cancer cells by inducing apoptosis through a receptor-mediated mechanism (Katayama *et al.,* 2006).

Proteinase K digested 46 kDa parasporal proteins of each of 5 indigenous Bt strains showed strong cytotoxic effects on the HeLa cell lines (Figure 5.3.10) than trypsin digested and crude protein. Round-shaped deceased HeLa carcinoma cells suggests that they probably died due to necrosis. In cancer cell killing parasporins the proteinase K digested PS2, PS4, and PS5 was found to induce necrosis-type cell death. Consequently, the cytotoxic effects of these strains on non-cancerous Vero cell lines were also investigated. Proteinase K digested parasporal proteins of 5 indigenous Bt strains didn't show any cytotoxic effects on the Vero cell lines.

As both parasporin (Mizuki *et al.*, 2000; Kim *et al.,* 2000; Nagamatsu *et al.*, 2010) and SLP (Rubio *et al.,* 2017; Zhang *et al.,* 2020, 2021) like anticancer activity was observed from the native parasporal proteins, so morphological examination of parasporal inclusions were conducted. The 0.2% Coomassie blue stained protein inclusions were found as somewhat round or oval or atypical shaped in the phase-contrast microscope. It was not conclusive from this observation that the inclusions were either Cry or SLP. To acquire a clear image about the source of parasporal protein, scanning electronic microscopy (SEM) was performed. From SEM, parasporal inclusions of all the strains were identified as round shaped with smooth surface which was in agreement with the literature where *B. thuringiensis* strain CTC and BMB171 demonstrated the same feature in SEM (Zhu & Yu, 2008). In contrast parasporin PS1Aa of Bt strain A1190 was reported as polygonal (Mizuki *et al.*, 2000). Besides, Cry1 was reported as bipyramidal, Cry2 as cuboidal, Cry3A as thin rhomboidal, Cry4 as atypical, Cry11A as bitrapezoidal,

Cyt1A as horse shoe shaped, and recombinant Cyt1A was reported as bihexagonal (Federici *et al.*, 2005). Hence, from the scanning electron microscopy (SEM) results of the parasporal inclusion proteins of the native Bt strains, it can be said that our identified parasporal anticancer proteins are of special type of parasporal protein. In light of this, the next obvious step ahead was to sequence the whole genomes of the strains in order to identify the proteins using bioinformatics.

### 7.1.4 Deciphering the genomic and proteomic profiles of indigenous *B. thuringiensis* strains with anticancer potential through WGS

To know the entire bacterial genome information, genotyping, and genetic features and any gene of interest, including their metabolic and biological subsystems, bacterial whole-genome sequencing (WGS) by next-generation sequencing (NGS) is an efficient and effective method. To identify the parasporal proteins isolated from five indigenous Bt isolates that demonstrated anticancer activity and to know their genotype, whole genome sequencing of four indigenous Bt strains BD59S, 28S, Soil 46, and 45L were conducted using the MiSeq System of Illumina NGS technology. The NGS-derived WGS data were assembled, annotated, and analyzed. Comparative studies were done as a requirement to prove bioinformatics data in the result section (Chapter 6) and a brief discussion is presented here for a better understanding.

#### 7.1.4.1 Analysis of the genomes

From the whole genome assembly, we identified the genome sizes of indigenous *B. thuringiensis* strains BD59S, Soil-46, 28S, and 45L as 5.28, 5.27, 5.27, and 5.37 megabases (Mb), respectively. It was found from the BRIG (Figure 6.3.4) and phylogenetic tree (Figure 6.3.6) that the genomes of BD59S, Soil-46 and 28S have a close relationship with the genome of Bt strain CTC (CP013274). The genome size of *B. thuringiensis* strain CTC is 5.32 MB (Dong *et al.*, 2016). Besides, 45L genome has a close relationship with the three Bt strain *konkukian* 97-27 (NC_005957), Bt 97-27 (CP010088), and BMBT15426 (CP020723). It was found that *B. thuringiensis* serovar *konkukian* str. 97-27 has a genome size of 5.23 MB (Han *et al.*, 2006), *B. thuringiensis* strain 97-27 has a genome size of 5.31 MB (Johnson *et al.*, 2016), and *B. thuringiensis* strain BM-BT15426 has a genome size of 5.24 MB (Liu *et al.*, 2017). Based on this

finding, it was apparent that the genome size of the indigenous Bt is directly related to the genome size of its closely related strain. From the phylogenetic tree and comparative genome analysis, it was determined that indigenous Bt strains BD59S, Soil-46, and 28S belonged to Bt serovar *finitimus*, and 45L belonged to the serovar *konkukian* (Figure 6.3.6).

One notable finding is that the Bt strain CTC was found to synthesize round parasporal inclusions, just like the parasporal inclusions of 5 native Bt strains. The parasporal inclusions produced in the Bt strain CTC were the S-layer proteins (Dong *et al.,* 2016). Here from comparative genome and phylogenetic tree analysis, it was also found that genome organization pattern (Figure 6.3.5), genome size, and phylogenetic relationships of indigenous Bt strains were very similar to the Bt strain CTC.

#### 7.1.4.2 Molecular analysis of the parasporal proteins

To determine the sequence of 101, 86, 76, and 103 kDa parasporal protein visualized on SDS-PAGE, WGS annotated genomic data of BD59S, 28S, Soil-46, and 45L were screened. No desired *cry* and *ps* gene were identified. Instead, four different S-layer protein and a hypothetical protein with desired molecular mass were identified.

Bt strains BD59S, 28S, Soil-46 produced 86 kDa and 101 kDa protein bands in the SDS-PAGE. From the WGS data we identified each of these three strains produced two S-layer protein, an 86 kDa Sap and a 91 kDa EA1. From the previous report, EA1 were visualized as 100 kDa protein in SDS-PAGE in many *B. thuringiensis* strains (Sun *et al.,* 2001; Peña *et al.,* 2006; Zhou *et al.,* 2011), while their actual molecular mass was 91 kDa. Considering that perspective, our identified 91 kDa EA1 of BD59S, 28S, Soil-46 could be the 101 kDa protein band in the SDS-PAGE.

In addition, round shaped parasporal proteins which were observed in the SEM are in agreement with the previous finding of Bt strain CTC (Sun *et al.,* 2001; Zhu & Yu, 2008). From this viewpoint, it may be assumed that the Sap and EA1 identified in these strains are the parasporal proteins.

From NCBI protein BLAST it was found that the 93.29 kDa SLP of 45L had 31% sequence identity with crystal protein of *B. thuringiensis* subsp. *mexicanensis* strain TKD2-14 (BAA13073.1) (Figure 6.3.7), and from the protein feature study, it was identified as an 823 amino acid long S-layer protein. This protein sequence was

submitted in the GenBank on 02-07-1996, by Naoki Hashimoto, Hokkaido University, Japan, and published in 07-02-1999. From their submission data we found that they submitted the sequence as Cry protein but it was a S-layer protein, that means this protein sequence also synthesized as parasporal inclusions in Bt. Besides, from NCBI protein BLAST, it was found that the 900 amino acids long 103.11 kDa S-layer protein of *B. thuringiensis* strains 45L had 99.56% identity with a protein sequence that was named as "parasporal protein" of *Streptococcus pneumoniae* (COE74891.1) (Figure 6.3.8). It was submitted by the Welcome Trust Sanger Institute of the United Kingdom but no additional information was obtained about this parasporal protein of *S. pneumoniae.* So, from these two BLASTs results it was found that both 93 kDa and 103 kDa SLP of 45L could be the parasporal SLP.

Multiple sequence alignment (MSA) of indigenous and other countries parasporal inclusion forming S-layer proteins (Figure 6.3.8) demonstrates the conserved SLH domain of the parasporal Sap, EA1, and other types of parasporal S-layer proteins differed significantly from each type, exhibiting a low degree of similarity (<30%). From MSA, it was found that parasporal SLPs are genealogically heterogeneous. The amino acid sequences of conserved surface layer homology domains (SLH) and the crystallization domains are highly diversified, whereas the amino acid sequences of signal peptides are homogeneous (Figure 6.3.8). Mizuki *et al.,* (2000) showed that parasporins and Cry proteins are genealogically heterogeneous which is also reflected in our findings.

From MAFFT alignments and the MView percent identity matrix, it was found that the 86 kDa S-layer protein Sap of 3 indigenous Bt strains had 96% to 85% amino acid sequence similarities with previously reported 86 kDa 4 parasporal S-layer protein Sap (ADU04486.1, ADQ08580.1, ADU04485.1, and CAA09981.1) of *B. thuringiensis* strains (Figure 6.3.9). Among these S-layer proteins, Sap, CAA09981.1 (85.57%), was reported to produce round parasporal inclusion bodies. Besides, 2 parasporal S-layer proteins, Sap ADQ08580.1 (96.04%), and ADU04485.1 (95.5%), were reported as ~50% identical with a recently published ~86 kDa MDA-MB-231 cervical cancer cell-killing parasporal S-layer protein (data not shown). Based on the MSA and the level of percent identity, it is reasonable to infer that there may be a strong correlation between

the 86 kDa SLP Sap and the production of parasporal inclusions in the indigenous Bt strains BD59S, 28S, and Soil-46.

The 91 kDa S-layer protein EA1 of 3 indigenous strains had 90% to 84% sequence similarities with previously reported 91 kDa 6 parasporal S-layer protein EA1 of different *B. thuringiensis* strains (Figure 6.3.10) (NCBI protein accession AAR23791.1, ADU04488.1, ADU04489.1, ADU04490.1, ADU04491.1, and AAY28601.1). Among these parasporal S-layer EA1, AAR23791.1 (87.54%) was reported to produce round parasporal inclusion bodies, and AAY28601.1 (84.17%) was reported to be toxic to the coleopteran beetle *Epilachna varivestis* (Peña *et al.,* 2006). Besides, indigenous 91 kDa parasporal S-layer EA1 protein had 76.59% identity with a recently published partial 66.26 kDa S-layer protein (protein sequencing data obtained from the publication) that was toxic to cattle tick *Rhipicephalus microplus* (Lormendez *et al.,* 2019). Based on the multiple sequence alignment (MSA) and the level of percent identity, it is reasonable to infer that there may be a strong correlation between the 91 kDa EA1 protein and the production of parasporal inclusions in the indigenous Bt strains BD59S, 28S, and Soil-46.

The 103 kDa SLP of indigenous Bt strain 45L had 16% sequence similarities with 6 parasporal SLPs of different Bt strains and it had maximum sequence similarities with EA1 (Figure 6.3.12). Besides, the 103 kDa SLP had 8% sequence similarities with the 93 kDa SLP of its own strain and 12% sequence similarities with the Sap of indigenous 3 Bt strains. It is worth mentioning that a significant segment, spanning from amino acids 28 to 309, of the N-terminal area of the 103 kDa SLP is a non-SLP region known as MABEL. This region is very different from the parasporal SLP EA1. Besides, from 708 to 901 amino acids is a cysteine peptidase enzyme region called YvpB, which also had somewhat similarities with the parasporal SLP EA1. Only the remaining 416 residue sequence is the real SLP region and hence these amino acids have very good similarities with other SLPs. But other than SLP regions, the large β-strand-rich protein regions contain the main crystallization domains; hence, only from the MSA and percent identity it was not clear that the relationships of the 103 kDa SLP with parasporal inclusion formation of Bt strain 45L. But, as its size was similar to the SDS-PAGE observed ~103 kDa parasporal protein bands, thus this protein could also be the parasporal inclusion forming SLP of the indigenous Bt strain 45L. So, it was very

difficult to identify from these two SLP which is the real parasporal inclusion forming protein.

#### 7.1.4.3 Molecular phylogeny, motif and domain analysis

A total of 21 S-layer proteins were studied for tree, motif and domain analysis. From the phylogenetic tree and motif study (Figure 6.3.17; Figure 6.3.18), it was clear that the 103 kDa SLP of indigenous strains 45L was largely different from other 20 studied S-layer proteins. The 93 kDa SLP of 45L was found closely related to the 87 kDa parasporal inclusion forming SLP of Bt mexicanensis, B22, and BMB1152. There was no evidence of what type of SLP they were, but from the phylogenetic tree and motif analysis, it was found that they had a closeness with the Sap. From comparative motif and phylogeny analysis between indigenous and experimentally proved parasporal inclusion-forming S-layer proteins of other Bt strains, it was believed that the 86 kDa Saps, the 91 kDa EA1s, and the 93 kDa S-layer proteins of indigenous Bt belonged to the parasporal inclusion-forming S-layer proteins of those strains. From comparative domain analysis it was very difficult to separate most of the SLPs according to their type or class (Figure 6.3.19). Hence, we found that to identify the relatedness between genealogical heterogenous or any other type of SLP, instead of domain analysis, motif and phylogenetic tree-based study is more effective. Finally, from this analysis, the 86 kDa Saps, the 91 kDa EA1s, and the 93 kDa S-layer protein of indigenous Bt strains BD59S, Soil-46, 28S, and 45L were found to be the most likely parasporal inclusions forming SLPs from those strains. Although, the 103 kDa SLP of 45L have identical in size to the SDS-PAGE observed parasporal protein, but, from comparative motif and phylogenetic tree analysis, it was thought the 93 kDa SLP of 45L is the indeed a round shaped parasporal inclusion protein.

#### 7.1.4.4 Functional domain analysis of the S-layer proteins

From the 91 kDa EA1 the β-strand-rich six β-sandwich crystallization or assembly domains were identified from the tertiary and secondary structures (Figure 6.3.22. A, C), a single amino acid change was found in the protein sequences of EA1 of BD59S. It was also found that trypsin-digested parasporal protein fragments of BD59S were different from those of Soil-46 and 28S because they didn’t produce 45 kDa and 32 kDa trypsin-digested protein fragments in BD59S, whereas Soil-46 and 28S did; hence,

bioinformatical analysis was conducted based on nucleotide alignment, protein's secondary structure comparison and 3D pdb structure alignment. As a result of single nucleotide polymorphism (Figure 6.3.23. A), the single amino acid was varied (Figure 6.3.23. B-C), and their distinct changes were observed both in the secondary and tertiary structures (Figure 6.3.23. B-C). From the secondary structure, it was found that EA1 of BD59S had a valine in 286 positions whereas a leucine was present in EA1 of Soil-46 and 28S. As a result, EA1 of BD59S had an alpha helix in 320 amino acid positions but instead an alpha helix in EA1 of Soil-46 and 28S had a loop (Figure 6.3.23. B). From the 3D tertiary structure alignment between these two EA1, it was clear that the position of the first β-strand-rich domain (after the 3 SLH domain) differentiated easily because they were apart from each other at 13.4 angstrom, and as a result, the 3 SLH domain also moved its orientation (Figure 6.3.23. C). This single amino acid difference in EA1 could be responsible for the origin of different types of tryptic digested protein fragments among these indigenous Bt strains.

The 93 kDa SLP of 45L had a 220 to 880 amino acid long β-strand-rich six β- sandwich crystallization or assembly domains (Figure 6.3.26. A). An aerolysin domain was identified from the SMART search tool (Figure 6.3.25) and this domain was found within the β-strand-rich crystallization domain, and about 3 β- sandwich regions belonged to this domain (Figure 6.3.26. A). Besides, two carbohydrate-binding modules, CBM2 (575 to 662), and CBM3 (770 to 836) were also found within the long β-strand-rich crystallization domain of the 93 kDa S-layer protein (Figure 6.3.26. A), and of them, CBM2 had sequence similarities with aerolysin and hence lay within the aerolysin domain. The aerolysin domains are a member of a major class of the β pore-forming toxins (β-PFTs). The family of aerolysin-like toxins damages cell membranes, which causes leaking and the eventual death of target cells (Cirauqui *et al.,* 2017). The β-PFT-type parasporins, are a type of aerolysin pore-forming toxins that can specifically kill various human cancer cell types by recognizing particular receptors on the plasma membrane of exposed cells (Akiba & Okumura, 2017). Besides, PS1, PS3, and PS6 had a CBM region at the C-terminal region. So, the β-strand-rich CBM2 and CBM3 domains of the 93 kDa SLP of 45L had a similarity with the parasporins.

The 103.11 kDa SLP had a ricin-type beta trefoil (661–770), and a YvpB cysteine peptidase C39-like domain (708–901) (Figure 6.3.28; Figure 6.3.29 A) in the long

crystallization region. Besides, 2 β-strand-rich domains were also observed in the 3D tertiary and 2D secondary protein structures (Figure 6.3.29, A, C). Ricin, which is a kind of protein that binds to carbohydrates, and it is a very powerful toxin that is produced in the seeds of the castor oil plant, *Ricinus communis.* Parasporin-3 was shown to possess a ricin domain, which contributes to the stabilization of the connection between toxins and carbohydrate residues on the cancer cell membrane (Krishnan *et al.,* 2017). In addition, the Cyt1Ca protein, which is produced by the pBtoxis plasmid of the Bt subsp. israelensis (Berry *et al.,* 2002), has a ricin carbohydrate-binding domain at the C-terminal end of the Cyt domain. However, no larvicidal or hemolytic activity has been detected with this toxin (Manasherob *et al.,* 2006). In addition, the Cry35Ab1 crystal structures that were recently published exhibit a ricin beta trefoil N-terminal domain that contains QxW motifs, comparable to the carbohydrate-binding domains observed in proteins like ricin and Mtx1 from *Lysinibacillus sphaericus* (Kelker *et al.,* 2014; Srisucharitpanit *et al.,* 2014). So, the 103 kDa SLP of 45L had a similar domain like parasporins-3, Cry and Cyt protein.

#### 7.1.4.5 Functional domain analysis of Cry-like proteins

Besides, the S-layer protein, a 616 amino acid long 73.54 kDa hypothetical protein molecule, was thought to be synthesized as a parasporal protein in Bt strains BD59S, Soil-46, and 28S, because its size was thought to be the same as the SDS PAGE observed ~76 kDa protein band. From InterPro scan, the NCBI protein features study, the NCBI conserved domain, and SMART search no domain is identified. From *in silico* tertiary and secondary protein structure prediction a three-domain Cry-like protein structure was formed from the 73.54 kDa protein sequences (Figure 6.3.24. A). Besides, from the automatic fold recognition and protein structure prediction psi-blast results of the PHYRE2 web server, it was found that the 73.54 kDa protein had 40% structure identity (from amino acids 126–140) with a tumor necrosis factor signaling protein region. Besides, it also had a 21-27% structure identity with two different types of DNA-binding proteins (from amino acids 112–198 and 489–582). Besides, 35.1% structure identity (from 479 to 575) with a nucleoplasmin-like/VP (viral coat and capsid proteins) of Picornaviridae-like positive stranded ssRNA viruses and 20% structure identity (415–458) with the hica3 toxin of *Yersinia pestis.* Besides, from the pdbsum webserver predicted secondary structure, it was distinctly found that the N-terminal and

the C-terminal portions of the protein molecule produced 2 β-strand rich domains, whereas the middle portion produced α-helices rich domains (Figure 6.3.24. C). So, from these observations, there could be some possibilities for the 73.54 kDa protein of these strains to be the Cry protein. Besides, the presence of this protein in all of three bacterium, and nearly same sizes parasporal protein band could be an additional indicator to identify it as a parasporal protein.

#### 7.1.4.6 Analysis of phylogenetic relationship of indigenous SLPs with parasporins and insecticidal Cry

The amino acid sequences of putative parasporal SLP, Cry-like 73.54 kDa protein, and six different classes of parasporins were used to make a phylogenetic tree. From the phylogenetic tree of indigenous SLPs, cry-like proteins, and six different classes of parasporins (Figure 6.3.29), it was found that SLPs had closeness with the PS2, PS4, and PS5. Besides, Cry-like 73.54 kDa protein had closeness with PS1, PS3, and PS6. Besides, SLPs and Cry-like 73.54 kDa proteins of indigenous Bt were also found to have some relationship with insecticidal Cry proteins. From the phylogenetic relationship and percent identity study (Figure 6.3.29; Figure 6.3.30; and 6.3.31 A-E), it was found that our identified SLPs and the 73.54 kDa protein have amino acid sequence similarities with the parasporal inclusion forming anticancer as well as some insecticidal Cry proteins. Cancer cell-killing Cry protein parasporins have very low homologies (<25%) with the existing classes of insecticidal Cry and Cyt proteins. We also found that our identified S-layer proteins (SLPs) had low homologies (<25%) with parasporins and some pesticidal Cry proteins.

As the anticancer activity was found in the proteinase K-digested and trypsin-digested parasporal proteins of five indigenous Bt strains, this analysis was performed to identify the probable anticancer parasporal protein. From WGS data analysis, four different types of S-layers (86 kDa Sap, 91 kDa EA1, 93 kDa, and 103 kDa) and 73.54 kDa cry-like proteins were identified as similar to the SDS-PAGE observed parasporal protein bands. Besides, from NCBI protein BLAST, multiple sequence alignments, percent identity, phylogenetic relationship, comparative motif, comparative domain, functional domain analysis of the S-layer and cry-like proteins of indigenous Bt strains, and finally their phylogenetic relationship and percentage of identity with anticancer parasporins and insecticidal cry, it can be said that our identified Sap, EA1, and the 93 kDa surface

layer protein are the most probable parasporal inclusion-forming proteins. Besides, the 103 kDa and the 73.54 kDa cry-like 3-domain proteins could also be parasporal inclusion-forming proteins, which need to be confirmed by expressing them in the acrystalliferous *B. thuringiensis* strains.

### 7.1.5 SWOT analysis

**Strength:** The strength of this study lies in the comprehensive screening and analysis of a large number of (121) *B. thuringiensis* (Bt) strains, leading to the exploration of parasporin, a rare anticancer protein. Primers for five well-known parasporin genes (ps1–ps5) were utilized to identify potential parasporin-encoding genes in both non-hemolytic and hemolytic strains. The extraction and purification of parasporal proteins revealed diverse molecular masses, some of which matched known parasporins. These proteins, once activated, exhibited selective cytotoxicity towards HeLa cancer cells while remaining non-toxic to normal Vero cells. Additionally, the proteins demonstrated an LC50 of 114.19 µg/mL against brine shrimp, were non-hemolytic to sheep red blood cells, and non-insecticidal to fruit fly larvae. WGS data confirmed the identity of specific protein bands as parasporin-like S-layer proteins, marking a novel discovery of their anticancer potential.

**Weakness:** Despite the substantial sample size of 121 Bt isolates, the genomic and proteomic findings did not match any known parasporins documented in existing databases, highlighting the rarity of parasporins in the environment and particularly in Bt, with only 19 identified to date. Additionally, the cytotoxicity of the putative parasporal proteins was not tested on a broader range of cancer cell lines due to the limited cell line availability during the study period. Further research needs to be carried out to address these gaps.

**Opportunity:** This study has opened several promising opportunities. Comprehensive genomic, proteomic, and bioinformatics analyses indicate that our country possesses a rich repertoire of indigenous *Bacillus* strains, which can be further inspected for the rare parasporin anticancer proteins. The cytotoxic effect of parasporal S-layer proteins from these indigenous Bt strains suggests their possible potential as candidates for developing targeted anticancer protein therapeutics. Further research could focus on identifying the specific anticancer domains and mechanisms of these proteins,

enhancing the development of effective protein-based therapeutics. Overall, the findings pave the way for more research into S-layer protein-derived anticancer proteins that could further be investigated for their potential as therapeutics in medical biotechnology using the tools and techniques that were not possible to explore in the present study.

**Threat:** Since the study was conducted in compliance with Good Laboratory Practice standards, no threats are anticipated from this study.

## 7.2 Conclusion

The study was aimed at isolating and characterizing parasporal anticancer proteins from indigenous *B. thuringiensis* strains. The round-shaped parasporal inclusion proteins from five Bt stains were found to be nonhemolytic and non-insecticidal and produced stable fragments after tryptic and/or proteinase K digestion, revealing a combination of parasporins and S-layer proteins with cytotoxic effects on HeLa cancer cells but not on Vero cells. These parasporal proteins were further identified and characterized using genomic, proteomic, and bioinformatics studies, revealing them as specific S-layer proteins and a three-domain Cry-like protein. Bioinformatics analysis further demonstrated similarities between parasporal S-layers and Cry proteins, including structural and functional domains. In this study, *B. thuringiensis* derived parasporal S-layer protein has been reported, which demonstrated anticancer activity against HeLa cancer cell lines without harming Vero cell lines, requiring further in-depth studies to uncover its anticancer potentials.

# CHAPTER 8:
# REFERENCES

# CHAPTER 8: REFERENCES


Abe, Y., Inoue, H., Ashida, H., Maeda, Y., Kinoshita, T., & Kitada, S. (2017). Glycan region of GPI anchored-protein is required for cytocidal oligomerization of an anticancer parasporin-2, Cry46Aa1 protein, from *Bacillus thuringiensis* strain A1547. *Journal of Invertebrate Pathology*, *142*, 71–81. https://doi.org/10.1016/j.jip.2016.11.008

Adang, M. J., Crickmore, N., & Jurat-Fuentes, J. L. (2014). *Diversity of Bacillus thuringiensis Crystal Toxins and Mechanism of Action*. https://doi.org/10.1016/B978-0-12-800197-4.00002-6

Agaisse, H., & Lereclus, D. (1995). How does *Bacillus thuringiensis* produce so much insecticidal crystal protein? *Journal of Bacteriology*, *177*(21), 6027–6032. https://doi.org/10.1128/jb.177.21.6027-6032.1995

Akiba, T., Abe, Y., Kitada, S., Kusaka, Y., Ito, A., Ichimatsu, T., … Harata, K. (2009). Crystal Structure of the Parasporin-2 *Bacillus thuringiensis* Toxin That Recognizes Cancer Cells. *Journal of Molecular Biology*, *386*(1), 121–133. https://doi.org/10.1016/j.jmb.2008.12.002

Akiba, T., & Okumura, S. (2017). Parasporins 1 and 2: Their structure and activity. *Journal of Invertebrate Pathology*, Vol. 142, pp. 44–49. Academic Press Inc. https://doi.org/10.1016/j.jip.2016.10.005

Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. *BMC Genomics*, *12*. https://doi.org/10.1186/1471-2164-12-402

Allievi, M. C., Palomino, M. M., Acosta, M. P., Lanati, L., Ruzal, S. M., & Sánchez-Rivas, C. (2014). Contribution of S-layer proteins to the mosquitocidal activity of Lysinibacillus sphaericus. *PLoS ONE*, *9*(10). https://doi.org/10.1371/journal.pone.0111114

Alp, D., Kuleaşan, H., & Korkut Altıntaş, A. (2020). The importance of the S-layer on the adhesion and aggregation ability of Lactic acid bacteria. *Molecular Biology Reports*, *47*(5), 3449–3457. https://doi.org/10.1007/s11033-020-05430-6

Al-yahyaee, S. A. S., & Ellar, D. J. (1995). Maximal toxicity of cloned CytA δ-endotoxin from *Bacillus thuringiensis* subsp. israelensis requires proteolytic processing from both the N- and C-termini. *Microbiology*, *141*(12), 3141–3148. https://doi.org/10.1099/13500872-141-12-3141

Andrews, R. E., Faust, R. M., Wabiko, H., Raymond, K. C., & Bulla, L. A. (1987). The biotechnology of *Bacillus thuringiensis. Critical Reviews in Biotechnology*, *6*(2), 163–232. https://doi.org/10.3109/07388558709113596

Andrews, S. (2010). *FASTQC. A quality control tool for high throughput sequence data.*

Angus, T. A. (1954). A Bacterial Toxin paralysing Silkworm Larvæ. *Nature*, *173*(4403), 545–546. https://doi.org/10.1038/173545a0

Aronson, A. I., Tyrell, D. J., Fitz-James, P. C., & Bulla, L. A. (1982). Relationship of the syntheses of spore coat protein and parasporal crystal protein in *Bacillus*

*thuringiensis*. *Journal of Bacteriology*, *151*(1), 399–410. https://doi.org/10.1128/jb.151.1.399-410.1982

Arrieta, G., Hernández, A., & Espinoza, A. M. (2004). Diversity of *Bacillus thuringiensis* strains isolated from coffee plantations infested with the coffee berry borer Hypothenemus hampei. *Revista de Biologia Tropical*, *52*(3), 757–764.

Asokan, R., Mahadeva Swamy, H. M., Thimmegowda, G. G., & Mahmood, R. (2014). Diversity analysis and characterization of Coleoptera-, Hemiptera- and Nematode-active cry genes in native isolates of *Bacillus thuringiensis*. *Annals of Microbiology*, *64*(1), 85–98. https://doi.org/10.1007/s13213-013-0636-7

Attathom, T., Chongrattanameteekul, W., Chanpaisang, J., & Siriyan, R. (1995). Morphological diversity and toxicity of delta-endotoxin produced by various strains of *Bacillus thuringiensis*. *Bulletin of Entomological Research*, *85*(2), 167–173. https://doi.org/10.1017/S0007485300034234

Aziz, R. K., Bartels, D., Best, A., DeJongh, M., Disz, T., Edwards, R. A., … Zagnitko, O. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*, *9*. https://doi.org/10.1186/1471-2164-9-75

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., … Noble, W. S. (2009). MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Research*, *37*(SUPPL. 2). https://doi.org/10.1093/nar/gkp335

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Barjac, H. de. (1981). Identification of H-serotypes of *Bacillus thuringiensis*. In B. HD (Ed.), *Microbial control of pests and plant diseases* (pp. 1970–1980). London Academic Press Inc. (London) Ltd.

Barjac, H., & Bonnefoi, A. (1962). Essai de classification biochimique et sérologique de 24 souches deBacillus du type *B. Thuringiensis*. *Entomophaga*, *7*(1), 5–31. https://doi.org/10.1007/BF02375988

Baumeister, W., & Engelhardt, H. (1987). *Three-dimensional Structure of Bacterial Surface Layers.* (Electron M; & R. W. Horne, J. Harris, Ed.). Retrieved from https://hdl.handle.net/11858/00-001M-0000-0010-74D8-8

Baumeister, W., Wildhaber, I., & Phipps, B. M. (1989). Principles of organization in eubacterial and archaebacterial surface proteins. *Canadian Journal of Microbiology*, *35*(1), 215–227. https://doi.org/10.1139/m89-034

Bechtel, D. B., & Bulla, L. A. (1976). Electron microscope study of sporulation and parasporal crystal formation in *Bacillus thuringiensis*. *Journal of Bacteriology*, *127*(3), 1472–1481. https://doi.org/10.1128/jb.127.3.1472-1481.1976

Bechtel, Donald B, & Bulla, L. A. (1982). Ultrastructural Analysis of Membrane Development during *Bacillus thuringiensis* Sporulation I. In *JOURNAL OF ULTRASTRUCTURE RESEARCH* (Vol. 79).

Bedini, S., Muniz, E. R., Tani, C., Conti, B., & Ruiu, L. (2020). Insecticidal potential of *Brevibacillus laterosporus* against dipteran pest species in a wide ecological

range. *Journal of Invertebrate Pathology*, *177*(October), 107493. https://doi.org/10.1016/j.jip.2020.107493

Ben-Dov, E. (2014). *Bacillus thuringiensis subsp. israelensis* and its dipteran-specific toxins. *Toxins*, *6*(4), 1222–1243. https://doi.org/10.3390/toxins6041222

Ben-Dov, E., Wang, Q., Zaritsky, A., Manasherob, R., Barak, Z., Schneider, B., … Margalith, Y. (1999). Multiplex PCR screening to detect cry9 genes in *Bacillus thuringiensis* strains. *Applied and Environmental Microbiology*, *65*(8), 3714–3716. https://doi.org/10.1128/aem.65.8.3714-3716.1999

Ben-dov, E., Zaritsky, A., Dahan, E., Barak, E., Sinai, R., Manasherob, R., … Margalith, Y. (1997). Extended Screening by PCR for Seven cry-Group Genes from Field-Collected Strains of *Bacillus thuringiensis*. In *APPLIED AND ENVIRONMENTAL MICROBIOLOGY* (Vol. 63).

Berliner, E. (1915). Berliner, E. (1915) Über die Schlaffsucht der Mehlmottehrauqe (*Ephestia kaechniella*), und invent Erreger *Bacillus thuringiensisn.* sp. Zeischrift fur angewandttes, Entomol., 2, 29-56. *Entomology*, (2), 29–56.

Bernhard, K., Jarrett, P., Meadows, M., Butt, J., Ellis, D. J., Roberts, G. M., … Burges, H. D. (1997). Natural Isolates of *Bacillus thuringiensis*:Worldwide Distribution, Characterization, and Activity against Insect Pests. *Journal of Invertebrate Pathology*, *70*(1), 59–68. https://doi.org/10.1006/jipa.1997.4669

Berry, C., O'Neil, S., Ben-Dov, E., Jones, A. F., Murphy, L., Quail, M. A., … Parkhill, J. (2002). Complete sequence and organization of pBtoxis, the toxin-coding plasmid of *Bacillus thuringiensis subsp. israelensis*. *Applied and Environmental Microbiology*, *68*(10), 5082–5095. https://doi.org/10.1128/AEM.68.10.5082-5095.2002

Bingle, W. H., Nomellini, J. F., & Smit, J. (1997). Linker Mutagenesis of the Caulobacter crescentus S-Layer Protein: toward a Definition of an N-Terminal Anchoring Region and a C-Terminal Secretion Signal and the Potential for Heterologous Protein Secretion. In *JOURNAL OF BACTERIOLOGY* (Vol. 179).

Bizzarri, M. F., & Bishop, A. H. (2008). The ecology of *Bacillus thuringiensis* on the Phylloplane: colonization from soil, plasmid transfer, and interaction with larvae of Pieris brassicae. *Microbial Ecology*, *56*(1), 133–139. https://doi.org/10.1007/s00248-007-9331-1

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boonserm, P., Davis, P., Ellar, D. J., & Li, J. (2005). Crystal structure of the mosquito-larvicidal toxin Cry4Ba and its biological implications. *Journal of Molecular Biology*, *348*(2), 363–382. https://doi.org/10.1016/j.jmb.2005.02.013

Boonserm, P., Mo, M., Angsuthanasombat, C., & Lescar, J. (2006). Structure of the functional form of the mosquito larvicidal Cry4Aa toxin from *Bacillus thuringiensis* at a 2.8-angstrom resolution. *Journal of Bacteriology*, *188*(9), 3391–3401. https://doi.org/10.1128/JB.188.9.3391-3401.2006

Boutonnet, C., Lyonnais, S., Alpha-Bazin, B., Armengaud, J., Château, A., & Duport, C. (2022). Dynamic Profile of S-Layer Proteins Controls Surface Properties of

Emetic *Bacillus cereus* AH187 Strain. *Frontiers in Microbiology*, *13*, 937862. https://doi.org/10.3389/fmicb.2022.937862

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, *72*, 248–254. https://doi.org/10.1006/abio.1976.9999

Brasseur, K., Auger, P., Asselin, E., Parent, S., Côté, J. C., & Sirois, M. (2015). Parasporin-2 from a new *Bacillus thuringiensis* 4r2 strain induces caspases activation and apoptosis in human cancer cells. *PLoS ONE*, *10*(8). https://doi.org/10.1371/journal.pone.0135106

Bravo, A, Sarabia, S., Lopez, L., Ontiveros, H., Abarca, C., Ortiz, A., … Quintero, R. (1998). Characterization of cry genes in a Mexican *Bacillus thuringiensis* strain collection. *Applied and Environmental Microbiology*, *64*(12), 4965–4972. https://doi.org/10.1128/AEM.64.12.4965-4972.1998

Bravo, Alejandra, Gómez, I., Porta, H., García-Gómez, B. I., Rodriguez-Almazan, C., Pardo, L., & Soberón, M. (2013). Evolution of *Bacillus thuringiensis* Cry toxins insecticidal activity. *Microbial Biotechnology*, *6*(1), 17–26. https://doi.org/10.1111/j.1751-7915.2012.00342.x

Bravo, Alejandra, Likitvivatanavong, S., Gill, S. S., & Soberón, M. (2011). *Bacillus thuringiensis*: A story of a successful bioinsecticide. *Insect Biochemistry and Molecular Biology*, *41*(7), 423–431. https://doi.org/10.1016/j.ibmb.2011.02.006

Brown, N. P., Leroy, C., & Sander, C. (1998). MView: a web-compatible database search or multiple alignment viewer. *Bioinformatics (Oxford, England)*, *14*(4), 380–381. https://doi.org/10.1093/bioinformatics/14.4.380

Bulla, L. A., Kramer, K. J., & Davidson, L. I. (1977). Characterization of the entomocidal parasporal crystal of *Bacillus thuringiensis*. *Journal of Bacteriology*, *130*(1), 375–383. https://doi.org/10.1128/jb.130.1.375-383.1977

Butko, P. (2003). Cytolytic toxin Cyt1A and its mechanism of membrane damage: Data and hypotheses. *Applied and Environmental Microbiology*, *69*(5), 2415–2422. https://doi.org/10.1128/AEM.69.5.2415-2422.2003

Candela, T., Mignot, T., Hagnerelle, X., Haustant, M., & Fouet, A. (2005). Genetic analysis of *Bacillus anthracis* Sap S-layer protein crystallization domain. *Microbiology*, *151*(5), 1485–1490. https://doi.org/10.1099/mic.0.27832-0

CANNON, R. J. C. (1996). *BACILLUS THURINGIENSIS* USE IN AGRICULTURE: A MOLECULAR PERSPECTIVE. *Biological Reviews*, *71*(4), 561–636. https://doi.org/10.1111/j.1469-185X.1996.tb01285.x

Carlson, C. R., & Kolstø, A. B. (1993). A complete physical map of a *Bacillus thuringiensis* chromosome. *Journal of Bacteriology*, *175*(4), 1053–1060. https://doi.org/10.1128/jb.175.4.1053-1060.1993

Carlton, B. C., and Gonzalez, J. M., Jr. (1984). The genetics and molecular biology of *Bacillus thuringiensis.* In D. Dubnau (Ed.), *The molecular biology of the bacilli* (pp. 211–249). New York: Acad. Press.

Carver, T., Thomson, N., Bleasby, A., Berriman, M., & Parkhill, J. (2009). DNAPlotter: Circular and linear interactive genome visualization. *Bioinformatics*, *25*(1), 119–120. https://doi.org/10.1093/bioinformatics/btn578

Ceron, J., Covarrubias, L., Quintero, R., Ortiz, A., Ortiz, M., Aranda, E., … Bravo, A. (1994). PCR analysis of the cryI insecticidal crystal family genes from *Bacillus thuringiensis*. *Applied and Environmental Microbiology*, *60*(1), 353–356. https://doi.org/10.1128/aem.60.1.353-356.1994

Ceron, J., Ortiz, A., Quintero, R., Guereca, L., & Bravo, A. (1995). Specific PCR primers directed to identify cryI and cryIII genes within a *Bacillus thuringiensis* strain collection. *Applied and Environmental Microbiology*, *61*(11), 3826–3831.

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., & Xia, R. (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular Plant*, *13*(8), 1194–1202. https://doi.org/10.1016/j.molp.2020.06.009

Chilcott, C. N., & Ellar, D. J. (1988). Comparative toxicity of *Bacillus thuringiensis var. israelensis* crystal proteins in vivo and in vitro. *Journal of General Microbiology*, *134*(9), 2551–2558. https://doi.org/10.1099/00221287-134-9-2551

Chilcott, C. N., & Wigley, P. J. (1993). Isolation and Toxicity of *Bacillus thuringiensis* from Soil and Insect Habitats in New Zealand. *Journal of Invertebrate Pathology*, *61*(3), 244–247. https://doi.org/10.1006/jipa.1993.1047

Chilcott, C. N., & Wigley, P. J. (1994). Opportunities for finding new *Bacillus thuringiensis* strains. *Agriculture, Ecosystems & Environment*, *49*(1), 51–57. https://doi.org/10.1016/0167-8809(94)90022-1

Choma, C. T., Surewicz, W. K., Carey, P. R., Pozsgay, M., Raynor, T., & Kaplan, H. (1990). Unusual proteolysis of the protoxin and toxin from *Bacillus thuringiensis*. Structural implications. *European Journal of Biochemistry*, *189*(3), 523–527. https://doi.org/10.1111/j.1432-1033.1990.tb15518.x

Choudhury, B., Leoff, C., Saile, E., Wilkins, P., Quinn, C. P., Kannenberg, E. L., & Carlson, R. W. (2006). The structure of the major cell wall polysaccharide of Bacillus anthracis is species-specific. *Journal of Biological Chemistry*, *281*(38), 27932–27941. https://doi.org/10.1074/jbc.M605768200

Cirauqui, N., Abriata, L. A., Van Der Goot, F. G., & Dal Peraro, M. (2017). Structural, physicochemical and dynamic features conserved within the aerolysin pore-forming toxin family. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-13714-4

Claus D, B. R. (1986). *Genus Bacillus Cohn 1872* (Bergey's m; H. J. Sneath PHA, Mair NS, Sharpe ME, Ed.). The Williams and Wilkins, Baltimore.

Claus, H., Akça, E., Debaerdemaeker, T., Evrard, C., Declercq, J.-P., & König, H. (2002). Primary Structure of Selected Archaeal Mesophilic and Extremely Thermophilic Outer Surface Layer Proteins. *Systematic and Applied Microbiology*, *25*(1), 3–12. https://doi.org/10.1078/0723-2020-00100

Cohen, S., Albeck, S., Ben-Dov, E., Cahan, R., Firer, M., Zaritsky, A., & Dym, O. (2011). Cyt1Aa toxin: Crystal structure reveals implications for its membrane-perforating function. *Journal of Molecular Biology*, *413*(4), 804–814. https://doi.org/10.1016/j.jmb.2011.09.021

Cohen, S., Dym, O., Albeck, S., Ben-Dov, E., Cahan, R., Firer, M., & Zaritsky, A. (2008). High-Resolution Crystal Structure of Activated Cyt2Ba Monomer from

*Bacillus thuringiensis subsp. israelensis. Journal of Molecular Biology*, *380*(5), 820–827. https://doi.org/10.1016/j.jmb.2008.05.010

Couture-Tosi, E., Delacroix, H., Mignot, T., Mesnage, S., Chami, M., Fouet, A., & Mosser, G. (2002). Structural analysis and evidence for dynamic emergence of *Bacillus anthracis* S-layer networks. *Journal of Bacteriology*, *184*(23), 6448–6456. https://doi.org/10.1128/JB.184.23.6448-6456.2002

Darling, A. C. E., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Research*, *14*(7), 1394–1403. https://doi.org/10.1101/gr.2289704

Darling, A. E., Mau, B., & Perna, N. T. (2010). Progressivemauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE*, *5*(6). https://doi.org/10.1371/journal.pone.0011147

Davidson, Elizabeth W., & Yamamoto, T. (1984). Isolation and assay of the toxic component from the crystals of *Bacillus thuringiensis* var. israelensis. *Current Microbiology*, *11*(3), 171–174. https://doi.org/10.1007/BF01567344

Davidson, Elizabeth West, & Myers, P. (1981). Parasporal inclusions in *Bacillus sphaericus. FEMS Microbiology Letters*, *10*(3), 261–265. Retrieved from https://www.sciencedirect.com/science/article/pii/0378109781902433

DeLucca, A. J., Simonson, J. G., & Larson, A. D. (1981). *Bacillus thuringiensis* distribution in soils of the United States. *Canadian Journal of Microbiology*, *27*(9), 865–870. https://doi.org/10.1139/m81-137

Dong, Z., Li, J., Zheng, J., Geng, C., Peng, D., & Sun, M. (2016). Complete genome sequence of *Bacillus thuringiensis* CTC-A typical strain with high production of S-layer proteins. *Journal of Biotechnology*, *220*, 100–101. https://doi.org/10.1016/j.jbiotec.2015.12.027

Drobniewski, F. A. (1993). *Bacillus cereus* and related species. *Clinical Microbiology Reviews*, *6*(4), 324–338. https://doi.org/10.1128/CMR.6.4.324

E. F. & SAMBROOK, J. M. T. F. (1982). *Molecular Cloning: a Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Egelseer, E. M., Leitner, K., Jarosch, M., Hotzy, C., Zayni, S., Sleytr, U. B., … Sa´ra, S. (1998). The S-Layer Proteins of Two Bacillus stearothermophilus Wild-Type Strains Are Bound via Their N-Terminal Region to a Secondary Cell Wall Polymer of Identical Chemical Composition. In *JOURNAL OF BACTERIOLOGY* (Vol. 180).

Ekino, K., Okumura, S., Ishikawa, T., Kitada, S., Saitoh, H., Akao, T., … Mizuki, E. (2014). Cloning and characterization of a unique cytotoxic protein parasporin-5 produced by *Bacillus thuringiensis* a1100 strain. *Toxins*, *6*(6), 1882–1895. https://doi.org/10.3390/toxins6061882

Engelhardt, Harald, & Peters, J. (1998). Structural research on surface layers: A focus on stability, surface layer homology domains, and surface layer-cell wall interactions. *Journal of Structural Biology*, *124*(2–3), 276–302. https://doi.org/10.1006/jsbi.1998.4070

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M.-Y., … Sali, A. (2006). Comparative protein structure modeling using Modeller.

*Current Protocols in Bioinformatics*, *Chapter 5*, Unit-5.6. https://doi.org/10.1002/0471250953.bi0506s15

Etienne-Toumelin, I., Sirard, J. C., Duflot, E., Mock, M., & Fouet, A. (1995). Characterization of the Bacillus anthracis S-layer: Cloning and sequencing of the structural gene. *Journal of Bacteriology*, *177*(3), 614–620. https://doi.org/10.1128/jb.177.3.614-620.1995

Federici, B. A. (2005). Insecticidal bacteria: An overwhelming success for invertebrate pathology. *Journal of Invertebrate Pathology*, *89*(1), 30–38. https://doi.org/10.1016/j.jip.2005.06.007

Fioravanti, A., Van Hauwermeiren, F., Van der Verren, S. E., Jonckheere, W., Goncalves, A., Pardon, E., … Remaut, H. (2019). Structure of S-layer protein Sap reveals a mechanism for therapeutic intervention in anthrax. *Nature Microbiology*, *4*(11), 1805–1814. https://doi.org/10.1038/s41564-019-0499-1

Frankenhuyzen, K. van. (2009, April). Insecticidal activity of *Bacillus thuringiensis* crystal proteins. *Journal of Invertebrate Pathology*, Vol. 101, pp. 1–16. https://doi.org/10.1016/j.jip.2009.02.009

Fujino, T., Beguin, P., & Aubert, J. P. (1993). Organization of a *Clostridium thermocellum* gene cluster encoding the cellulosomal scaffolding protein CipA and a protein possibly involved in attachment of the cellulosome to the cell surface. *Journal of Bacteriology*, *175*(7), 1891–1899. https://doi.org/10.1128/jb.175.7.1891-1899.1993

Galitsky, N., Cody, V., Wojtczak, A., Ghosh, D., Luft, J. R., Pangborn, W., & English, L. (2001). Structure of the insecticidal bacterial delta-endotoxin Cry3Bb1 of *Bacillus thuringiensis*. *Acta Crystallographica. Section D, Biological Crystallography*, *57*(Pt 8), 1101–1109. https://doi.org/10.1107/s0907444901008186

Gill, S. S., Singh, G. J., & Hornung, J. M. (1987). Cell membrane interaction of *Bacillus thuringiensis subsp. israelensis* cytolytic toxins. *Infection and Immunity*, *55*(5), 1300–1308. https://doi.org/10.1128/iai.55.5.1300-1308.1987

Gleave, A. P., Hedges, R. J., & Broadwell, A. H. (1992). Identification of an insecticidal crystal protein from *Bacillus thuringiensis* DSIR517 with significant sequence differences from previously described toxins. In *Journal of General Microbiology* (Vol. 138).

Goldberg, L.J., & Margalit, J. (1977). A bacterial spore demonstrating rapid larvicidal activity against *Anopheles sergentii, Uranotaenia unguiculata, Culex univittatus, Aedes aegypti* and *Culex pipiens*. *Mosquito News*, *37*(3), 355–358.

Gonzalez, E., Granados, J. C., Short, J. D., Ammons, D. R., & Rampersad, J. (2011). Parasporins from a Caribbean Island: evidence for a globally dispersed *Bacillus thuringiensis* strain. *Current Microbiology*, *62*(5), 1643–1648. https://doi.org/10.1007/s00284-011-9905-5

Grisolia, C. K., Oliveira-Filho, E. C., Ramos, F. R., Lopes, M. C., Muniz, D. H. F., & Monnerat, R. G. (2009). Acute toxicity and cytotoxicity of *Bacillus thuringiensis* and Bacillus sphaericus strains on fish and mouse bone marrow. *Ecotoxicology*, *18*(1), 22–26. https://doi.org/10.1007/s10646-008-0252-7

Grochulski, P., Masson, L., Borisova, S., Pusztai-Carey, M., Schwartz, J.-L., Brousseau, R., & Cygler, M. (1995). *Bacillus thuringiensis* CrylA(a) Insecticidal Toxin: Crystal Structure and Channel Formation. *Journal of Molecular Biology*, *254*(3), 447–464. https://doi.org/10.1006/jmbi.1995.0630

Guillet, P., Kurtak, D. C., Philippon, B., & Meyer, R. (1990). Use of *Bacillus thuringiensis* israelensis for Onchocerciasis Control in West Africa. In *Bacterial Control of Mosquitoes & Black Flies* (pp. 187–201). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-011-5967-8_11

Guo, G., Zhang, L., Zhou, Z., Ma, Q., Liu, J., Zhu, C., … Sun, M. (2008). A new group of parasporal inclusions encoded by the S-layer gene of *Bacillus thuringiensis*. *FEMS Microbiology Letters*, *282*(1), 1–7. https://doi.org/10.1111/j.1574-6968.2008.01087.x

Guo, S., Ye, S., Liu, Y., Wei, L., Xue, J., Wu, H., … Rao, Z. (2009). Crystal structure of *Bacillus thuringiensis* Cry8Ea1: An insecticidal toxin toxic to underground pests, the larvae of Holotrichia parallela. *Journal of Structural Biology*, *168*(2), 259–266. https://doi.org/10.1016/j.jsb.2009.07.004

Guo, Y., Weng, M., Sun, Y., Carballar-Lejarazú, R., Wu, S., & Lian, C. (2022). *Bacillus thuringiensis* toxins with nematocidal activity against the pinewood nematode Bursaphelenchus xylophilus. *Journal of Invertebrate Pathology*, *189*, 107726. https://doi.org/10.1016/j.jip.2022.107726

Guttmann, D. M., & Ellar, D. J. (2000). Phenotypic and genotypic comparisons of 23 strains from the Bacillus cereus complex for a selection of known and putative *B. thuringiensis* virulence factors. *FEMS Microbiology Letters*, *188*(1), 7–13. https://doi.org/10.1016/S0378-1097(00)00200-7

Han, C. S., Xie, G., Challacombe, J. F., Altherr, M. R., Bhotika, S. S., Bruce, D., … Gilna, P. (2006). Pathogenomic sequence analysis of *Bacillus cereus* and *Bacillus thuringiensis* isolates closely related to *Bacillus anthracis*. *Journal of Bacteriology*, *188*(9). https://doi.org/10.1128/JB.188.9.3382-3390.2006

Hannay, C. L., & Fitz-James, P. (1955). THE PROTEIN CRYSTALS OF *BACILLUS THURINGIENSIS* BERLINER. *Canadian Journal of Microbiology*, *1*(8), 694–710. https://doi.org/10.1139/m55-083

Hannay, Christopher L. (1961). *FOWLER'S BACILLUS AND ITS PARASPORAL BODY*.

Hansen, B. M., Damgaard, P. H., Eilenberg, J., & Pedersen, J. C. (1998). Molecular and Phenotypic Characterization of *Bacillus thuringiensis* Isolated from Leaves and Insects. *Journal of Invertebrate Pathology*, *71*(2), 106–114. https://doi.org/10.1006/jipa.1997.4712

Hastowo, S., Lay, B. W., & Ohba, M. (1992). Naturally occurring *Bacillus thuringiensis* in Indonesia. *Journal of Applied Bacteriology*, *73*(2), 108–113. https://doi.org/10.1111/j.1365-2672.1992.tb01695.x

Heimpel, A. M. (1967). A taxonomic key proposed for the species of the "crystalliferous bacteria". *Journal of Invertebrate Pathology*, *9*(3), 364–375. https://doi.org/10.1016/0022-2011(67)90072-9

Heimpel, A. M., & Aangus, T. A. (1958). The taxonomy of insect pathogens related to Bacillus cereus Frankland and Frankland. *Canadian Journal of Microbiology*, *4*(5), 531–541. https://doi.org/10.1139/m58-058

Helgason, E., Caugant, D. A., Lecadet, M. M., Chen, Y., Mahillon, J., Lövgren, A., … Kolstø, A. B. (1998). Genetic diversity of *Bacillus cereus/B. thuringiensis* isolates from natural sources. *Current Microbiology*, *37*(2), 80–87. https://doi.org/10.1007/s002849900343

Hendriksen, N. B., & Hansen, B. M. (2002). Long-term survival and germination of *Bacillus thuringiensis* var. kurstaki in a field trial. *Canadian Journal of Microbiology*, *48*(3), 256–261. https://doi.org/10.1139/w02-009

Höfte, H., & Whiteley, H. R. (1989). Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiological Reviews*, *53*(2), 242–255. https://doi.org/10.1128/mr.53.2.242-255.1989

Howorka, S., Sára, M., Wang, Y., Kuen, B., Sleytr, U. B., Lubitz, W., & Bayley, H. (2000). Surface-accessible residues in the monomeric and assembled forms of a bacterial surface layer protein. *Journal of Biological Chemistry*, *275*(48), 37876–37886. https://doi.org/10.1074/jbc.M003838200

https://saves.mbi.ucla.edu/. (n.d.).

https://www.bioinformatics.org/sms/prot_mw.html. (n.d.).

Ichikawa, M., Uemori, A., Yasutake, K., Kagoshima, K., Mizuki, E., & Ohba, M. (2008). Failure to phenotypically discriminate between non-insecticidal *Bacillus thuringiensis* strains with anticancer parasporins (PS2, PS3, and PS4) and *Bacillus thuringiensis* strains that produce insecticidal Cry proteins. *Applied Entomology and Zoology*, *43*(3), 421–426. https://doi.org/10.1303/aez.2008.421

Ichimatsu, T., Mizuki, E., Nishimura, K., Akao, T., Saitoh, H., Higuchi, K., & Ohba, M. (2000). Occurrence of *Bacillus thuringiensis* in Fresh Waters of Japan. *Current Microbiology*, *40*(4), 217–220. https://doi.org/10.1007/s002849910044

Ilk, N., Kosma, P., Puchberger, M., Egelseer, E. M., Mayer, H. F., Sleytr, U. B., & Sára, M. (1999). Structural and functional analyses of the secondary cell wall polymer of *Bacillus sphaericus* CCM 2177 that serves as an S-layer-specific anchor. *Journal of Bacteriology*, *181*(24), 7643–7646. https://doi.org/10.1128/JB.181.24.7643-7646.1999

Ishii, T., & Ohba, M. (1993). Diversity of *Bacillus thuringiensis* environmental isolates showing larvicidal activity specific for mosquitoes. *Journal of General Microbiology*, *139*(11), 2849–2854. https://doi.org/10.1099/00221287-139-11-2849

Ishiwatari, S. (1901). On a kind of severe flacherie (sotto disease). *Dainihon Sanshi Kaiho*, (114), 1–5.

Ito, A., Sasaguri, Y., Kitada, S., Kusaka, Y., Kuwano, K., Masutomi, K., … Ohba, M. (2004). A *Bacillus thuringiensis* crystal protein with selective cytocidal action to human cells. *Journal of Biological Chemistry*, *279*(20), 21282–21286. https://doi.org/10.1074/jbc.M401881200

Jack, D. L., Yang, N. M., & Saier, M. H. (2001). The drug/metabolite transporter superfamily. *European Journal of Biochemistry*, *268*(13), 3620–3639. https://doi.org/10.1046/j.1432-1327.2001.02265.x

Jallouli, W., Driss, F., Fillaudeau, L., & Rouis, S. (2020). Review on biopesticide production by *Bacillus thuringiensis* subsp. kurstaki since 1990: Focus on bioprocess parameters. *Process Biochemistry*, *98*, 224–232. https://doi.org/10.1016/j.procbio.2020.07.023

Jarosch, M., Egelseer, E. M., Huber, C., Moll, D., Mattanovich, D., Sleytr, U. B., & Sára, M. (2001). Analysis of the structure–function relationship of the S-layer protein SbsC of *Bacillus stearothermophilus* ATCC 12980 by producing truncated forms. *Microbiology*, *147*(5), 1353–1363. https://doi.org/10.1099/00221287-147-5-1353

Jarosch, M., Egelseer, E. M., Mattanovich, D., Sleytr, U. B., Sa, M., & Ra, ! (2000). S-layer gene sbsC of Bacillus stearothermophilus ATCC 12980 : molecular characterization and heterologous expression in Escherichia coli. In *Microbiology* (Vol. 146).

Jing, H., Takagi, J., Liu, J., Lindgren, S., Zhang, R., Joachimiak, A., … Springer, T. A. (2002). Archaeal surface layer proteins contain beta propeller, PKD, and beta helix domains and are related to metazoan cell surface proteins. *Structure (London, England : 1993)*, *10*(10), 1453–1464. https://doi.org/10.1016/s0969-2126(02)00840-7

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Research*, *36*(Web Server issue). https://doi.org/10.1093/nar/gkn201

Johnson, S. L., Daligault, H. E., Davenport, K. W., Jaissle, J., Frey, K. G., Ladner, J. T., … Chain, P. S. (2016). Complete genome sequences for 35 biothreat assay-relevant *Bacillus* species. *Genome Announcements*, *3*(2). https://doi.org/10.1128/genomeA.00151-15

Jouzani, G. S., Abad, A. P., Seifinejad, A., Marzban, R., Kariman, K., & Maleki, B. (2008). Distribution and diversity of Dipteran-specific cry and cyt genes in native *Bacillus thuringiensis* strains obtained from different ecosystems of Iran. *Journal of Industrial Microbiology & Biotechnology*, *35*(2), 83–94. https://doi.org/10.1007/s10295-007-0269-6

Jouzani, G. S., Valijanian, E., & Sharafi, R. (2017). *Bacillus thuringiensis*: a successful insecticide with new environmental features and tidings. *Applied Microbiology and Biotechnology*, Vol. 101, pp. 2691–2711. Springer Verlag. https://doi.org/10.1007/s00253-017-8175-y

Jung, Y. C., Mizuki, E., Akao, T., & Côté, J. C. (2007). Isolation and characterization of a novel *Bacillus thuringiensis* strain expressing a novel crystal protein with cytocidal activity against human cancer cells. *Journal of Applied Microbiology*, *103*(1), 65–79. https://doi.org/10.1111/j.1365-2672.2006.03260.x

Kaelin, P., Morel, P., & Gadani, F. (1994). Isolation of *Bacillus thuringiensis* from Stored Tobacco and Lasioderma serricorne (F.). *Applied and Environmental Microbiology*, *60*(1), 19–25. https://doi.org/10.1128/aem.60.1.19-25.1994

Kalendar, R., Khassenov, B., Ramankulov, Y., Samuilova, O., & Ivanov, K. I. (2017). FastPCR: An *in silico* tool for fast primer and probe design and advanced sequence analysis. *Genomics*, *109*(3–4), 312–319. https://doi.org/10.1016/j.ygeno.2017.05.005

Kalendar, R., Lee, D., & Schulman, A. H. (2014). FastPCR software for PCR, in silico PCR, and oligonucleotide assembly and analysis. *Methods in Molecular Biology (Clifton, N.J.)*, *1116*, 271–302. https://doi.org/10.1007/978-1-62703-764-8_18

Katayama H, Kusaka Y, M. E. (2011). *Patent No. EP2273266A1*. European Patent Office.

Katayama, H., Kusaka, Y., Yokota, H., Akao, T., Kojima, M., Nakamura, O., … Mizuki, E. (2007). Parasporin-1, a novel cytotoxic protein from *Bacillus thuringiensis*, induces Ca2+ influx and a sustained elevation of the cytoplasmic Ca2+ concentration in toxin-sensitive cells. *Journal of Biological Chemistry*, *282*(10), 7742–7752. https://doi.org/10.1074/jbc.M611382200

Katoh, K., Rozewicki, J., & Yamada, K. D. (2018). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, *20*(4), 1160–1166. https://doi.org/10.1093/bib/bbx108

Kelker, M. S., Berry, C., Evans, S. L., Pai, R., McCaskill, D. G., Wang, N. X., … Narva, K. E. (2014). Structural and biophysical characterization of *Bacillus thuringiensis* insecticidal proteins Cry34Ab1 and Cry35Ab1. *PloS One*, *9*(11), e112555. https://doi.org/10.1371/journal.pone.0112555

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, *10*(6), 845–858. https://doi.org/10.1038/nprot.2015.053

Kern, J., Wilton, R., Zhang, R., Binkowski, T. A., Joachimiak, A., & Schneewind, O. (2011). Structure of surface layer homology (SLH) domains from *Bacillus anthracis* surface array protein. *Journal of Biological Chemistry*, *286*(29), 26042–26049. https://doi.org/10.1074/jbc.M111.248070

Kim, D. E., Chivian, D., & Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research*, *32*(Web Server issue), W526-31. https://doi.org/10.1093/nar/gkh468

Kim, H. S., Yamashita, S., Akao, T., Saitoh, H., Higuchi, K., Park, Y. S., … Ohba, M. (2000). In vitro cytotoxicity of non-Cyt inclusion proteins of a *Bacillus thuringiensis* isolate against human cells, including cancer cells. *Journal of Applied Microbiology*, *89*(1), 16–23. https://doi.org/10.1046/j.1365-2672.2000.01087.x

Kitada, S., Abe, Y., Maeda, T., & Shimada, H. (2009). Parasporin-2 requires GPI-anchored proteins for the efficient cytocidal action to human hepatoma cells. *Toxicology*, *264*(1–2), 80–88. https://doi.org/10.1016/j.tox.2009.07.016

Kitada, S., Abe, Y., Shimada, H., Kusaka, Y., Matsuo, Y., Katayama, H., … Ito, A. (2006). Cytocidal actions of parasporin-2, an anti-tumor crystal toxin from *Bacillus thuringiensis*. *Journal of Biological Chemistry*, *281*(36), 26350–26360. https://doi.org/10.1074/jbc.M602589200

Knowles, B. H. (1994). *Mechanism of Action of Bacillus thuringiensis Insecticidal δ-Endotoxins*. https://doi.org/10.1016/S0065-2806(08)60085-5

Knowles, B. H., Blatt, M. R., Tester, M., Horsnell, J. M., Carroll, J., Menestrina, G., & Ellar, D. J. (1989). A cytolytic δ-endotoxin from *Bacillus thuringiensis* var. israelensis forms cation-selective channels in planar lipid bilayers. *FEBS Letters*, *244*(2), 259–262. https://doi.org/10.1016/0014-5793(89)80540-X

Koller, C. N., Bauer, L. S., & Hollingworth, R. M. (1992). Characterization of the pH-mediated solubility of *Bacillus thuringiensis* var. san diego native delta-endotoxin crystals. *Biochemical and Biophysical Research Communications*, *184*(2), 692–699. https://doi.org/10.1016/0006-291x(92)90645-2

Kondo, S., Mizuki, E., Akao, T., & Ohba, M. (2002). Antitrichomonal strains of *Bacillus thuringiensis*. *Parasitology Research*, *88*(12), 1090–1092. https://doi.org/10.1007/s00436-002-0692-6

Konit, P. A., & Ellar, D. J. (1994). Biochemical characterization of *Bacillus thuringiensis* cytolytic δ endotoxins. *Microbiology*, *140*, 1869-1 880.

Krieg, A., Huger, A. M., Langenbruch, G. A., & Schnetter, W. (1983). *Bacillus thuringiensis* var. tenebrionis : ein neuer, gegenüber Larven von Coleopteren wirksamer Pathotyp. *Zeitschrift Für Angewandte Entomologie*, *96*(1–5), 500–508. https://doi.org/10.1111/j.1439-0418.1983.tb03704.x

Krieg, A., Schnetter, W., Huger, A. M., & Langenbruch, G. A. (1987). *Bacillus thuringiensis* subsp. *tenebrionis*, strain BI 256–82: a third pathotype within the H-serotype 8a8b. *Systematic and Applied Microbiology*, *9*(1–2), 138–141. https://doi.org/10.1016/S0723-2020(87)80068-1

Krishnan, K., Ker, J. E. A., Mohammed, S. M., & Nadarajah, V. D. (2010). Identification of Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as a binding protein for a 68-kDa *Bacillus thuringiensis* parasporal protein cytotoxic against leukaemic cells. *Journal of Biomedical Science*, *17*(1), 1–11. https://doi.org/10.1186/1423-0127-17-86

Krishnan, V., Domanska, B., Elhigazi, A., Afolabi, F., West, M. J., & Crickmore, N. (2017). The human cancer cell active toxin Cry41Aa from *Bacillus thuringiensis* acts like its insecticidal counterparts. *Biochemical Journal*, *474*(10). https://doi.org/10.1042/BCJ20170122

Kumar, P., Kamle, M., Borah, R., Mahato, D. K., & Sharma, B. (2021). *Bacillus thuringiensis* as microbial biopesticide: uses and application for sustainable agriculture. *Egyptian Journal of Biological Pest Control*, *31*(1), 95. https://doi.org/10.1186/s41938-021-00440-3

Kuroda, S., Begum, A., Saga, M., Hirao, A., Mizuki, E., Sakai, H., & Hayakawa, T. (2013). Parasporin 1Ac2, a Novel Cytotoxic Crystal Protein Isolated from *Bacillus thuringiensis* B0462 Strain. *Current Microbiology*, *66*(5), 475–480. https://doi.org/10.1007/s00284-013-0301-1

Lambert, B., & Peferoen, M. (1992). Insecticidal Promise of *Bacillus thuringiensis*. *BioScience*, *42*(2), 112–122. https://doi.org/10.2307/1311652

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S., & Thornton, J. M. (2018). PDBsum: Structural summaries of PDB entries. *Protein Science: A Publication of the Protein Society*, *27*(1), 129–134. https://doi.org/10.1002/pro.3289

Lecadet, M.-M., Frachon, E., Dumanoir, V. C., Ripouteau, H., Hamon, S., Laurent, P., & Thié, I. (1999). Updating the H-antigen classification of *Bacillus thuringiensis*. In *Journal of Applied Microbiology* (Vol. 86).

Lee, D. W., Akao, T., Yamashita, S., Katayama, H., Maeda, M., Saitoh, H., … Ohba, M. (2000). Noninsecticidal parasporal proteins of a *Bacillus thuringiensis* serovar shandongiensis isolate exhibit a preferential cytotoxicity against human leukemic T cells. *Biochemical and Biophysical Research Communications*, *272*(1), 218–223. https://doi.org/10.1006/bbrc.2000.2765

Legler, D. F., Doucey, M.-A., Schneider, P., Chapatte, L., Bender, F. C., & Bron, C. (2005). Differential insertion of GPI-anchored GFPs into lipid rafts of live cells. *The FASEB Journal*, *19*(1), 73–75. https://doi.org/10.1096/fj.03-1338fje

Lenina, N. K., Naveenkumar, A., Sozhavendan, A. E., Balakrishnan, N., Balasubramani, V., & Udayasuriyan, V. (2014). Characterization of parasporin gene harboring Indian isolates of *Bacillus thuringiensis. 3 Biotech*, *4*(5), 545–551. https://doi.org/10.1007/s13205-013-0190-9

Letunic, I., Khedkar, S., & Bork, P. (2021). SMART: recent updates, new developments and status in 2020. *Nucleic Acids Research*, *49*(D1), D458–D460. https://doi.org/10.1093/nar/gkaa937

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, J. D., Carroll, J., & Ellar, D. J. (1991). Crystal structure of insecticidal delta-endotoxin from *Bacillus thuringiensis* at 2.5 A resolution. *Nature*, *353*(6347), 815–821. https://doi.org/10.1038/353815a0

Li, J., Koni, P. A., & Ellar, D. J. (1996). Structure of the mosquitocidal delta-endotoxin CytB from *Bacillus thuringiensis* sp. kyushuensis and implications for membrane pore formation. *Journal of Molecular Biology*, *257*(1), 129–152. https://doi.org/10.1006/jmbi.1996.0152

Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., … Lopez, R. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Research*, *43*(W1), W580–W584. https://doi.org/10.1093/nar/gkv279

Lingwood, C. A., Manis, A., Mahfoud, R., Khan, F., Binnington, B., & Mylvaganam, M. (2010). New aspects of the regulation of glycosphingolipid receptor function. *Chemistry and Physics of Lipids*, *163*(1), 27–35. https://doi.org/10.1016/j.chemphyslip.2009.09.001

Liu, C., Zheng, J., Ou, X., & Han, Y. (2021). Anti-cancer Substances and Safety of Lactic Acid Bacteria in Clinical Treatment. *Frontiers in Microbiology*, *12*(October). https://doi.org/10.3389/fmicb.2021.722052

Liu, J., Li, L., Peters, B. M., Li, B., Chen, D., Xu, Z., & Shirtliff, M. E. (2017). Complete genome sequence and bioinformatics analyses of *Bacillus thuringiensis* strain

BM-BT15426. *Microbial Pathogenesis*, *108*. https://doi.org/10.1016/j.micpath.2017.05.006

Lormendez, C. C., Fernandez-Ruvalcaba, M., Adames-Mancebo, M., Hernandez-Velazquez, V. M., Zuñiga-Navarrete, F., Flores-Ramirez, G., … Peña-Chora, G. (2019). Mass production of a S-layer protein of *Bacillus thuringiensis* and its toxicity to the cattle tick *Rhipicephalus microplus*. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-53854-3

Lupas, A. (1996). A circular permutation event in the evolution of the SLH domain? *Molecular Microbiology*, *20*(4), 897–898. https://doi.org/10.1111/j.1365-2958.1996.tb02528.x

Lupas, Andrei, Engelhardt, H., Peters, J., Santarius, U., Volker, S., & Baumeister, W. (1994). Domain Structure of the *Acetogenium kivui* Surface Layer Revealed by Electron Crystallography and Sequence Analysis. In *JOURNAL OF BACTERIOLOGY* (Vol. 176).

Lynch, M. J., & Baumann, P. (1985). Immunological comparisons of the crystal protein from strains of *Bacillus thuringiensis. Journal of Invertebrate Pathology*, *46*(1), 47–57. https://doi.org/10.1016/0022-2011(85)90128-4

Mader, C., Huber, C., Moll, D., Sleytr, U. B., & Sára, M. (2004). Interaction of the Crystalline Bacterial Cell Surface Layer Protein SbsB and the Secondary Cell Wall Polymer of *Geobacillus stearothermophilus* PV72 Assessed by Real-Time Surface Plasmon Resonance Biosensor Technology. *Journal of Bacteriology*, *186*(6), 1758–1768. https://doi.org/10.1128/JB.186.6.1758-1768.2004

Mahillon, J., Rezsöhazy, R., Hallet, B., & Delcour, J. (1994). IS 231 and other *Bacillus thuringiensis* transposable elements: A review. *Genetica*, *93*(1–3), 13–26. https://doi.org/10.1007/BF01435236

Manasherob, R., Itsko, M., Sela-Baranes, N., Ben-Dov, E., Berry, C., Cohen, S., & Zaritsky, A. (2006). Cyt1Ca from *Bacillus thuringiensis subsp. israelensis*: production in Escherichia coli and comparison of its biological activities with those of other Cyt-like proteins. *Microbiology (Reading, England)*, *152*(Pt 9), 2651–2659. https://doi.org/10.1099/mic.0.28981-0

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., … Bryant, S. H. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Research*, *43*(Database issue), D222-6. https://doi.org/10.1093/nar/gku1221

Marshall, O. J. (2004). PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics (Oxford, England)*, *20*(15), 2471–2472. https://doi.org/10.1093/bioinformatics/bth254

May, A., Pusztahelyi, T., Hoffmann, N., Fischer, R.-J., & Bahl, H. (2006). Mutagenesis of conserved charged amino acids in SLH domains of *Thermoanaerobacterium thermosulfurigenes* EM1 affects attachment to cell wall sacculi. *Archives of Microbiology*, *185*(4), 263–269. https://doi.org/10.1007/s00203-006-0092-x

Meadows, M. P., Ellis, D. J., Butt, J., Jarrett, P., & Denis Burges, H. (1992). Distribution, Frequency, and Diversity of *Bacillus thuringiensis* in an Animal Feed Mill. In *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*.

Mesnage, S, Weber-Levy, M., Haustant, M., Mock, M., & Fouet, A. (1999). Cell surface-exposed tetanus toxin fragment C produced by recombinant *Bacillus*

*anthracis* protects against tetanus toxin. *Infection and Immunity*, *67*(9), 4847–4850. https://doi.org/10.1128/IAI.67.9.4847-4850.1999

Mesnage, Stéphane, Fontaine, T., Mignot, T., Delepierre, M., Mock, M., & Fouet, A. (2000). Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *EMBO Journal*, *19*(17), 4473–4484. https://doi.org/10.1093/emboj/19.17.4473

Mignot, T., Mesnage, S., Couture-Tosi, E., Mock, M., & Fouet, A. (2002). Developmental switch of S-layer protein synthesis in *Bacillus anthracis*. *Molecular Microbiology*, *43*(6), 1615–1627. https://doi.org/10.1046/j.1365-2958.2002.02852.x

Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., & Marshall, D. (2009). Tablet-next generation sequence assembly visualization. *Bioinformatics*, *26*(3), 401–402. https://doi.org/10.1093/bioinformatics/btp666

Mizuki, E, Ohba, M., Akao, T., Yamashita, S., Saitoh, H., & Park, Y. S. (1999). Unique activity associated with non-insecticidal *Bacillus thuringiensis* parasporal inclusions: in vitro cell-killing action on human cancer cells. In *Journal of Applied Microbiology* (Vol. 86).

Mizuki, E, Park, Y. S., Saitoh, H., Yamashita, S., Akao, T., Higuchi, K., & Ohba, M. (2000). Parasporin, a human leukemic cell-recognizing parasporal protein of *Bacillus thuringiensis. Clinical and Diagnostic Laboratory Immunology*, *7*(4), 625–634. https://doi.org/10.1128/CDLI.7.4.625-634.2000

Montaldi, F. A., & Roth, I. L. (1990). Parasporal bodies of *Bacillus laterosporus* sporangia. *Journal of Bacteriology*, *172*(4), 2168–2171. https://doi.org/10.1128/jb.172.4.2168-2171.1990

Morse, R. J., Yamamoto, T., & Stroud, R. M. (2001). Structure of Cry2Aa suggests an unexpected receptor binding epitope. *Structure (London, England : 1993)*, *9*(5), 409–417. https://doi.org/10.1016/s0969-2126(01)00601-3

Mounsef, J. R., Salameh, D., Awad, M. kallassy, Chamy, L., Brandam, C., & Lteif, R. (2014). A simple method for the separation of *Bacillus thuringiensis* spores and crystals. *Journal of Microbiological Methods*, *107*, 147–149. https://doi.org/10.1016/j.mimet.2014.10.003

Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., … Yeats, C. (2007). New developments in the InterPro database. *Nucleic Acids Research*, *35*(SUPPL. 1). https://doi.org/10.1093/nar/gkl841

Nagamatsu, Y., Okamura, S., Saitou, H., Akao, T., & Mizuki, E. (2010). Three cry toxins in two types from *Bacillus thuringiensis* strain M019 preferentially kill human hepatocyte cancer and uterus cervix cancer cells. *Bioscience, Biotechnology and Biochemistry*, *74*(3), 494–498. https://doi.org/10.1271/bbb.90615

Nicolopoulou-Stamati, P., Maipas, S., Kotampasi, C., Stamatis, P., & Hens, L. (2016). Chemical Pesticides and Human Health: The Urgent Need for a New Concept in Agriculture. *Frontiers in Public Health*, *4*. https://doi.org/10.3389/fpubh.2016.00148

Ohba, M. (1996). *Bacillus thuringiensis* populations naturally occurring on mulberry leaves: a possible source of the populations associated with silkworm-rearing

insectaries. *Journal of Applied Bacteriology*, *80*(1), 56–64. https://doi.org/10.1111/j.1365-2672.1996.tb03190.x

Ohba, M., Yu, Y. M., & Aizawa, K. (1988). Occurrence of Non-Insecticidal *Bacillus thuringiensis* Flagellar Serotype 14 in the Soil of Japan. *Systematic and Applied Microbiology*, *11*(1), 85–89. https://doi.org/10.1016/S0723-2020(88)80052-3

Ohba, Michio, & Aizawa, K. (1986). Insect toxicity of *Bacillus thuringiensis* isolated from soils of Japan. *Journal of Invertebrate Pathology*, *47*(1), 12–20. https://doi.org/10.1016/0022-2011(86)90158-8

Ohba, Michio, Mizuki, E., & Uemori, A. (2009). Parasporin, a new anticancer protein group from *Bacillus thuringiensis*. *Anticancer Research*, *29*(1), 427–433.

Okumura, S., Ishikawa, T., Saitoh, H., Akao, T., & Mizuki, E. (2013). Identification of a second cytotoxic protein produced by *Bacillus thuringiensis* A1470. *Biotechnology Letters*, *35*(11), 1889–1894. https://doi.org/10.1007/s10529-013-1275-6

Okumura, S., Koga, H., Inouye, K., & Mizuki, E. (2014). Toxicity of Parasporin-4 and Health Effects of Pro-parasporin-4 Diet in Mice. *Toxins*, *6*(7), 2115–2126. https://doi.org/10.3390/toxins6072115

Okumura, S., Saitoh, H., Ishikawa, T., Inouye, K., & Mizuki, E. (2011). Mode of action of parasporin-4, a cytocidal protein from *Bacillus thuringiensis*. *Biochimica et Biophysica Acta - Biomembranes*, *1808*(6), 1476–1482. https://doi.org/10.1016/j.bbamem.2010.11.003

Okumura, S., Saitoh, H., Ishikawa, T., Mizuki, E., & Inouye, K. (2008). Identification and characterization of a novel cytotoxic protein, parasporin-4, produced by *Bacillus thuringiensis* A1470 strain. *Biotechnology Annual Review*, Vol. 14, pp. 225–252. https://doi.org/10.1016/S1387-2656(08)00009-4

Olson, R. D., Assaf, R., Brettin, T., Conrad, N., Cucinell, C., Davis, J. J., … Stevens, R. L. (2023). Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. *Nucleic Acids Research*, *51*(1 D), D678–D689. https://doi.org/10.1093/nar/gkac1003

Palma, L., Muñoz, D., Berry, C., Murillo, J., & Caballero, P. (2014). *Bacillus thuringiensis* toxins: An overview of their biocidal activity. *Toxins*, Vol. 6, pp. 3296–3325. MDPI. https://doi.org/10.3390/toxins6123296

Patel, K. D., Chudasama, C. J., & Ingle, S. S. (2011). Molecular characterization of *Bacillus thuringiensis* isolated from diverse habitats of India. *Journal of Basic Microbiology*, *52*(4), 437–445. https://doi.org/10.1002/jobm.201100080

Pavkov, T., Oberer, M., Egelseer, E. M., Sára, M., Sleytr, U. B., & Keller, W. (2003). Crystallization and preliminary structure determination of the C-terminal truncated domain of the S-layer protein SbsC. *Acta Crystallographica Section D Biological Crystallography*, *59*(8), 1466–1468. https://doi.org/10.1107/S0907444903010990

Payne, J. M., & Davidson, E. W. (1984). Insecticidal activity of the crystalline parasporal inclusions and other components of the *Bacillus sphaericus* 1593 spore complex. *Journal of Invertebrate Pathology*, *43*(3), 383–388. https://doi.org/10.1016/0022-2011(84)90084-3

Peña, G., Miranda-Rios, J., De La Riva, G., Pardo-López, L., Soberón, M., & Bravo, A. (2006). A *Bacillus thuringiensis* S-layer protein involved in toxicity against *Epilachna varivestis* (Coleoptera: Coccinellidae). *Applied and Environmental Microbiology*, *72*(1), 353–360. https://doi.org/10.1128/AEM.72.1.353-360.2006

Petkau, A., Stuart-Edwards, M., Stothard, P., & van Domselaar, G. (2010). Interactive microbial genome visualization with GView. *Bioinformatics*, *26*(24), 3125–3126. https://doi.org/10.1093/bioinformatics/btq588

Ponting, C. P., Schultz, J., Milpetz, F., & Bork, P. (1999). SMART: identification and annotation of domains from signalling and extracellular protein sequences. *Nucleic Acids Research*, *27*(1), 229–232. https://doi.org/10.1093/nar/27.1.229

Poornima, K., Selvanayagam, P., & Shenbagarathai, R. (2010). Identification of native *Bacillus thuringiensis* strain from South India having specific cytocidal activity against cancer cells. *Journal of Applied Microbiology*, *109*(1), 348–354. https://doi.org/10.1111/j.1365-2672.2010.04697.x

Poppinga, L., Janesch, B., Fünfhaus, A., Sekot, G., Garcia-Gonzalez, E., Hertlein, G., … Genersch, E. (2012). Identification and functional analysis of the S-layer protein SplA of *Paenibacillus larvae*, the causative agent of American Foulbrood of honey bees. *PLoS Pathogens*, *8*(5). https://doi.org/10.1371/journal.ppat.1002716

Porcar, M., & Juárez-Pérez, V. (2003). PCR-based identification of *Bacillus thuringiensis* pesticidal crystal genes. *FEMS Microbiology Reviews*, *26*(5), 419–432. https://doi.org/10.1111/j.1574-6976.2003.tb00624.x

Pozsgay, M., Fast, P., Kaplan, H., & Carey, P. R. (1987). The effect of sunlight on the protein crystals from *Bacillus thuringiensis* var. kurstaki HD1 and NRD12: A Raman spectroscopic study. *Journal of Invertebrate Pathology*, *50*(3), 246–253. https://doi.org/10.1016/0022-2011(87)90089-9

Prasadi, S. S. S. V, & Shethna, Y. I. (1976). Mode of Action of a Purified Antitumor Protein from the Proteinaceous Crystal of *Bacillus thuringiensis* subsp. thuringiensis on Yoshida Ascites Sarcoma Cells. In *ANTIMICROBIAL AGENTS AND CHZMYrHZRAPY*.

Promdonky, B., & Ellar, D. J. (2003). Investigation of the pore-forming mechanism of a cytolytic δ-endotoxin from *Bacillus thuringiensis*. *Biochemical Journal*, *374*(1), 255–259. https://doi.org/10.1042/bj20030437

Pum, D., Toca-Herrera, J. L., & Sleytr, U. B. (2013). S-Layer protein self-assembly. *International Journal of Molecular Sciences*, Vol. 14, pp. 2484–2501. https://doi.org/10.3390/ijms14022484

Pundir, S., Martin, M. J., & O'Donovan, C. (2016). UniProt Tools. *Current Protocols in Bioinformatics*, *53*(1), 1.29.1-1.29.15. https://doi.org/10.1002/0471250953.bi0129s53

Rambaut, A. (2010). *FigTree v1.4.0. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh. http://tree.bio.ed.ac.uk/software/figtree/*.

Rampersad, J., Khan, A., & Ammons, D. (2002). Usefulness of staining parasporal bodies when screening for *Bacillus thuringiensis*. *Journal of Invertebrate Pathology*, *79*(3), 203–204. https://doi.org/10.1016/S0022-2011(02)00018-6

Rasko, D. A., Altherr, M. R., Han, C. S., & Ravel, J. (2005). Genomics of the *Bacillus cereus* group of organisms. *FEMS Microbiology Reviews*, *29*(2), 303–329. https://doi.org/10.1016/j.femsre.2004.12.005

Reyes-ramírez, A., Escudero-Abarca, B. I., Aguilar-Uscanga, G., Hayward-Jones, P. M., & Barboza-Corona, J. E. (2004). Antifungal Activity of *Bacillus thuringiensis* Chitinase and Its Potential for the Biocontrol of Phytopathogenic Fungi in Soybean Seeds. *Journal of Food Science*, *69*(5). https://doi.org/10.1111/j.1365-2621.2004.tb10721.x

Rigden, D. J. (2009). Does distant homology with Evf reveal a lipid binding site in *Bacillus thuringiensis* cytolytic toxins? *FEBS Letters*, *583*(10), 1555–1560. https://doi.org/10.1016/j.febslet.2009.04.038

Rowe, G. E., Margaritis, A., & Dulmage, H. T. (1987). *BIOPROCESS DEVELOPMENTS IN THE PRODUCTION OF BIOINSECTICIDES BY BACILLUS THURZNGZENSZS* (Vol. 6).

Ruan, L., Crickmore, N., Peng, D., & Sun, M. (2015). Are nematodes a missing link in the confounded ecology of the entomopathogen *Bacillus thuringiensis*? *Trends in Microbiology*, *23*(6), 341–346. https://doi.org/10.1016/j.tim.2015.02.011

Rubikas, J., Androsiuniene, D., Chestukhina, G., Smirnova, T., Kapitonova, O., & Stepanov, V. (1987). Crystal protein formed by *Bacillus subtilis* cells. *Journal of Bacteriology*, *169*(11), 5258–5262. https://doi.org/10.1128/jb.169.11.5258-5262.1987

Rubio, V. P., Bravo, A., & Olmos, J. (2017). Identification of a *Bacillus thuringiensis* surface layer protein with cytotoxic activity against MDA-MB-231 breast cancer cells. *Journal of Microbiology and Biotechnology*, *27*(1). https://doi.org/10.4014/jmb.1607.07020

Ruiu, L. (2023). Multiple S-Layer Proteins of *Brevibacillus laterosporus* as Virulence Factors against Insects. *International Journal of Molecular Sciences*, *24*(2). https://doi.org/10.3390/ijms24021781

Rünzler, D., Huber, C., Moll, D., Köhler, G., & Sára, M. (2004). Biophysical Characterization of the Entire Bacterial Surface Layer Protein SbsB and Its Two Distinct Functional Domains. *Journal of Biological Chemistry*, *279*(7), 5207–5215. https://doi.org/10.1074/jbc.M308819200

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., & Barrell, B. (2000). Artemis: sequence visualization and annotation. *Bioinformatics (Oxford, England)*, *16*(10), 944–945. https://doi.org/10.1093/bioinformatics/16.10.944

Sambrook, J., Fritsch, E. R., Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual* (2nd ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Santos, E. N., Menezes, L. P., Dolabella, S. S., Santini, A., Severino, P., Capasso, R., … Jain, S. (2022). *Bacillus thuringiensis*: From biopesticides to anticancer agents. *Biochimie*, *192*, 83–90. https://doi.org/10.1016/j.biochi.2021.10.003

Sára, M., Kuen, B., Mayer, H. F., Mandl, F., Schuster, K. C., & Sleytr, U. B. (1996). Dynamics in oxygen-induced changes in S-layer protein synthesis from *Bacillus stearothermophilus* PV72 and the S-layer-deficient variant T5 in continuous

culture and studies of the cell wall composition. *Journal of Bacteriology*, *178*(7), 2108–2117. https://doi.org/10.1128/jb.178.7.2108-2117.1996

Sára, M., & Sleytr, U. B. (2000). S-Layer Proteins. *Journal of Bacteriology*, *182*(4), 859–868. https://doi.org/10.1128/JB.182.4.859-868.2000

Sawaya, M. R., Cascio, D., Gingery, M., Rodriguez, J., Goldschmidt, L., Colletier, J.-P., … Eisenberg, D. S. (2014). Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(35), 12769–12774. https://doi.org/10.1073/pnas.1413456111

Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., … Dean, D. H. (1998). *Bacillus thuringiensis* and Its Pesticidal Crystal Proteins. In *MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS* (Vol. 62).

Schultz, J., Milpetz, F., Bork, P., & Ponting, C. P. (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences*, *95*(11), 5857–5864. https://doi.org/10.1073/pnas.95.11.5857

Shishir, A., Roy, A., Islam, N., Rahman, A., Khan, S. N., & Hoq, M. M. (2014). Abundance and diversity of *Bacillus thuringiensis* in Bangladesh and their cry genes profile. *Frontiers in Environmental Science*, *2*(JUN). https://doi.org/10.3389/fenvs.2014.00020

Sievers, F., & Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sequences. *Protein Science*, *27*(1), 135–145. https://doi.org/10.1002/pro.3290

Sillanpää, J., Martínez, B., Antikainen, J., Toba, T., Kalkkinen, N., Tankka, S., … Korhonen, T. K. (2000). Characterization of the collagen-binding S-layer protein CbsA of *Lactobacillus crispatus*. *Journal of Bacteriology*, *182*(22), 6440–6450. https://doi.org/10.1128/JB.182.22.6440-6450.2000

Slepecky, R. A., & Hemphill, H. E. (2006). The Genus Bacillus—Nonmedical. In E. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, KH., Stackebrandt (Ed.), *The Prokaryotes*. https://doi.org/10.1007/0-387-30744-3_16

Sleytr, U B, Messner, P., Pum, D., & Sára, M. (1993). Crystalline bacterial cell surface layers. *Molecular Microbiology*, *10*(5), 911–916. https://doi.org/10.1111/j.1365-2958.1993.tb00962.x

Sleytr, Uwe B., & Beveridge, T. J. (1999). Bacterial S-layers. *Trends in Microbiology*, *7*(6), 253–260. https://doi.org/10.1016/S0966-842X(99)01513-9

Sleytr, Uwe B., Messner, P., Pum, D., & Sára, M. (1993). Crystalline bacterial cell surface layers. *Molecular Microbiology*, *10*(5), 911–916. https://doi.org/10.1111/j.1365-2958.1993.tb00962.x

Sleytr, Uwe B., Schuster, B., Egelseer, E. M., & Pum, D. (2014). S-layers: Principles and applications. *FEMS Microbiology Reviews*, *38*(5), 823–864. https://doi.org/10.1111/1574-6976.12063

Sleytr, Uwe B, & Sfira, M. (1986). *Mini-review from crystalline bacterial cell envelope layers*. 83–90.

Smirnova, T. A., Minenkova, I. B., Orlova, M. V., Lecadet, M. M., & Azizbekyan, R. R. (1996). The crystal-forming strains of *Bacillus laterosporus*. *Research in Microbiology*, *147*(5), 343–350. https://doi.org/10.1016/0923-2508(96)84709-7

Smit, E., Jager, D., Martinez, B., Tielen, F. J., & Pouwels, P. H. (2002). Structural and functional analysis of the S-layer protein crystallisation domain of *Lactobacillus acidophilus* ATCC 4356: Evidence for protein-protein interaction of two subdomains. *Journal of Molecular Biology*, *324*(5), 953–964. https://doi.org/10.1016/S0022-2836(02)01135-X

Smit, E., Oling, F., Demel, R., Martinez, B., & Pouwels, P. H. (2001). The S-layer protein of *Lactobacillus acidophilus* ATCC 4356: Identification and characterisation of domains responsible for S-protein assembly and cell wall binding. *Journal of Molecular Biology*, *305*(2), 245–257. https://doi.org/10.1006/jmbi.2000.4258

Smith, R. A., & Couche, G. A. (1991). The Phylloplane as a Source of *Bacillus thuringiensis* Variants. *Applied and Environmental Microbiology*, *57*(1), 311–315. https://doi.org/10.1128/aem.57.1.311-315.1991

Soberón, M., López-Díaz, J. A., & Bravo, A. (2013). Cyt toxins produced by *Bacillus thuringiensis*: a protein fold conserved in several pathogenic microorganisms. *Peptides*, *41*, 87–93. https://doi.org/10.1016/j.peptides.2012.05.023

Soufiane, B., & Côté, J. C. (2009). Discrimination among *Bacillus thuringiensis* H serotypes, serovars and strains based on 16S rRNA, gyrB and aroE gene sequence analyses. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *95*(1), 33–45. https://doi.org/10.1007/s10482-008-9285-4

Soufiane, B., Sirois, M., & Côté, J. C. (2011). Mutually exclusive distribution of the sap and eag S-layer genes and the lytB/lytA cell wall hydrolase genes in *Bacillus thuringiensis. Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *100*(3), 349–364. https://doi.org/10.1007/s10482-011-9590-1

Srisucharitpanit, K., Yao, M., Promdonkoy, B., Chimnaronk, S., Tanaka, I., & Boonserm, P. (2014). Crystal structure of BinB: A receptor binding component of the binary toxin from Lysinibacillus sphaericus. *Proteins: Structure, Function, and Bioinformatics*, *82*(10), 2703–2712. https://doi.org/10.1002/prot.24636

Stahly, D. P., Andrews, R. E., & Yousten, A. A. (2006). The Genus Bacillus—Insect Pathogens. In *The Prokaryotes* (pp. 563–608). New York, NY: Springer US. https://doi.org/10.1007/0-387-30744-3_17

Steinhaus, E. A. (1951). Possible Use of *Bacillus thuringiensis* Berliner as an Aid in the Biological Control of the Alfalfa Caterpillar. *Hilgardia*, *20*(18), 359–381.

Sun, M., Zhu, C., & Yu, Z. (2001). Cloning of parasporal body protein gene resembling to S-layer protein genes from *Bacillus thuringiensis* CTC strain. *Wei Sheng Wu Xue Bao = Acta Microbiologica Sinica*, *41*(2), 141–147.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A*

*Cancer Journal for Clinicians*, *71*(3), 209–249. https://doi.org/10.3322/caac.21660

Swamy, H. M. M., & Asokan, R. (2013). New Innovations in *Bacillus thuringiensis* Research- A Review. *International Journal of Agriculture, Environment and Biotechnology*, *6*(4), 545. https://doi.org/10.5958/j.2230-732x.6.4.029

Swiecicka, I., Fiedoruk, K., & Bednarz, G. (2002). The occurrence and properties of *Bacillus thuringiensis* isolated from free-living animals. *Letters in Applied Microbiology*, *34*(3), 194–198. https://doi.org/10.1046/j.1472-765x.2002.01070.x

Sychantha, D., Chapman, R. N., Bamford, N. C., Boons, G. J., Howell, P. L., & Clarke, A. J. (2018). Molecular Basis for the Attachment of S-Layer Proteins to the Cell Wall of *Bacillus anthracis*. *Biochemistry*, *57*(13), 1949–1953. https://doi.org/10.1021/acs.biochem.8b00060

Tabashnik, B. E., Zhang, M., Fabrick, J. A., Wu, Y., Gao, M., Huang, F., ... & Li, X. (2015). Dual mode of action of Bt proteins: protoxin efficacy against resistant insects. Scientific reports, 5(1), 15107. https://doi.org/10.1038/srep15107

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, *30*(12), 2725–2729. https://doi.org/10.1093/molbev/mst197

Tatusova, T., Dicuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., … Ostell, J. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research*, *44*(14), 6614–6624. https://doi.org/10.1093/nar/gkw569

Thanabalu, T., Hindley, J., Jackson-Yap, J., & Berry, C. (1991). Cloning, sequencing, and expression of a gene encoding a 100-kilodalton mosquitocidal toxin from *Bacillus sphaericus* SSII-1. *Journal of Bacteriology*, *173*(9), 2776–2785. https://doi.org/10.1128/jb.173.9.2776-2785.1991

*The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.* (n.d.).

Thomas, W. E., & Ellar, D. J. (1983). *Bacillus thuringiensis var israelensis* crystal delta-endotoxin: effects on insect and mammalian cells in vitro and in vivo. *Journal of Cell Science*, *60*, 181–197. https://doi.org/10.1242/jcs.60.1.181

Tojo, A., & Aizawa, K. (1983). Dissolution and Degradation of *Bacillus thuringiensis* delta-Endotoxin by Gut Juice Protease of the Silkworm Bombyx mori. *Applied and Environmental Microbiology*, *45*(2), 576–580. https://doi.org/10.1128/aem.45.2.576-580.1983

Travers, R. S., Martin, P. A. W., & Reichelderfer, C. F. (1987). Selective Process for Efficient Isolation of Soil Bacillus spp. In *APPLIED AND ENVIRONMENTAL MICROBIOLOGY* (Vol. 53).

Travis R. Glare, M. O. (2000). *Bacillus thuringiensis: Biology, Ecology and Safety*. John Wiley & Sons, Chichester.

Tyrell, D. J., Bulla, L. A., Andrews, R. E., Kramer, K. J., Davidson, L. I., & Nordin, P. (1981). Comparative biochemistry of entomocidal parasporal crystals of selected *Bacillus thuringiensis* strains. *Journal of Bacteriology*, *145*(2), 1052–1062. https://doi.org/10.1128/jb.145.2.1052-1062.1981

Uemori, A., Maeda, M., Yasutake, K., Ohgushi, A., Kagoshima, K., Mizuki, E., & Ohba, M. (2007). Ubiquity of parasporin-1 producers in *Bacillus thuringiensis*

natural populations of Japan. *Die Naturwissenschaften*, *94*(1), 34–38. https://doi.org/10.1007/s00114-006-0153-7

Uemori, A., Ohgushi, A., Yasutake, K., Maeda, M., Mizuki, E., & Ohba, M. (2008). Parasporin-1Ab, a novel *Bacillus thuringiensis* cytotoxin preferentially active on human cancer cells in vitro. *Anticancer Research*, *28*(1A), 91–95.

van Frankenhuyzen, K. (2013). Cross-order and cross-phylum activity of *Bacillus thuringiensis* pesticidal proteins. *Journal of Invertebrate Pathology*, *114*(1), 76–85. https://doi.org/10.1016/j.jip.2013.05.010

Walters, F. S., DeFontes, C. M., Hart, H., Warren, G. W., & Chen, J. S. (2010). Lepidopteran-Active Variable-Region Sequence Imparts Coleopteran Activity in eCry3.1Ab, an Engineered *Bacillus thuringiensis* Hybrid Insecticidal Protein. *Applied and Environmental Microbiology*, *76*(10), 3082–3088. https://doi.org/10.1128/AEM.00155-10

Wang, F., Liu, Y., Zhang, F., Chai, L., Ruan, L., Peng, D., … Valentino, L. (1987). Selective Process for Efficient Isolation of Soil Bacillus spp. *Applied and Environmental Microbiology*, *53*(20), 7197–7204. https://doi.org/10.1128/AEM.01048-12

Wasano, N., Saitoh, H., Maeda, M., Ohgushi, A., Mizuki, E., & Ohba, M. (2005). Cloning and characterization of a novel gene cry9Ec1 encoding lepidopteran-specific parasporal inclusion protein from a *Bacillus thuringiensis* serovar galleriae strain. *Canadian Journal of Microbiology*, *51*(11), 988–995. https://doi.org/10.1139/w05-084

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., … Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303. https://doi.org/10.1093/nar/gky427

Wong, S. Y. R. (2010). 2.Review_Rebecca. *International E-Journal of Science, Medicine and Education*, *4*(2), 3–9.

Wong., S. Y. R. (2010). *Bacillus thuringiensis* Parasporal Proteins And Their Effect On Human Cancer Cells : An Overview. *International E-Journal of Science, Medicine and Education*, *4*(2), 3–9. Retrieved from https://pesquisa.bvsalud.org/portal/resource/pt/wpr-629313

Wu, D., & Chang, F. N. (1985). Synergism in mosquitocidal activity of 26 and 65 kDa proteins from *Bacillus thuringiensis* subsp. israelensis crystal. *FEBS Letters*, *190*(2), 232–236. https://doi.org/10.1016/0014-5793(85)81290-4

Wu, M., Liu, X., Bai, H., Lai, L., Chen, Q., Huang, G., … Tang, G. (2019). Surface-Layer Protein-Enhanced Immunotherapy Based on Cell Membrane-Coated Nanoparticles for the Effective Inhibition of Tumor Growth and Metastasis [Research-article]. *ACS Applied Materials and Interfaces*, *11*(10), 9850–9859. https://doi.org/10.1021/acsami.9b00294

Yamamoto, T., & McLaughlin, R. E. (1981). Isolation of a protein from the parasporal crystal of *Bacillus thuringiensis* var, kurstaki toxic to the mosquito larva, Aedes taeniorhynchus. *Biochemical and Biophysical Research Communications*, *103*(2), 414–421. https://doi.org/10.1016/0006-291X(81)90468-X

Yamashita, S., Akao, T., Mizuki, E., Saitoh, H., Higuchi, K., Shin Park, Y., … Ohba, M. (2000). Characterization of the anti-cancer-cell parasporal proteins of a *Bacillus thuringiensis* isolate. *Canadian Journal of Microbiology*, *46*(10), 913–919. https://doi.org/10.1139/w00-084

Yamashita, Satoko, Katayama, H., Saitoh, H., Akao, T., Park, Y. S., Mizuki, E., … Ito, A. (2005). Typical three-domain cry proteins of *Bacillus thuringiensis* strain A1462 exhibit cytocidal activity on limited human cancer cells. *Journal of Biochemistry*, *138*(6), 663–672. https://doi.org/10.1093/jb/mvi177

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nature Methods*, *12*(1), 7–8. https://doi.org/10.1038/nmeth.3213

Yasutake, K., Binh, N. D., Kagoshima, K., Uemori, A., Ohgushi, A., Maeda, M., … Ohba, M. (2006). Occurrence of parasporin-producing *Bacillus thuringiensis* in Vietnam. *Canadian Journal of Microbiology*, *52*(4), 365–372. https://doi.org/10.1139/w05-134

Zaru, R., & Orchard, S. (2023). UniProt Tools: BLAST, Align, Peptide Search, and ID Mapping. *Current Protocols*, *3*(3), 1–26. https://doi.org/10.1002/cpz1.697

Zelazny, B., Stephan, D., & Hamacher, J. (1994). Irregular Crystal Formation in Some Isolates of *Bacillus thuringiensis*. *Journal of Invertebrate Pathology*, *63*(3), 229–234. https://doi.org/10.1006/jipa.1994.1045

Zhang, T., Pan, D., Yang, Y., Jiang, X., Zhang, J., Zeng, X., … Guo, Y. (2020). Effect of Lactobacillus acidophilus CICC 6074 S-Layer Protein on Colon Cancer HT-29 Cell Proliferation and Apoptosis. *Journal of Agricultural and Food Chemistry*, *68*(9), 2639–2647. https://doi.org/10.1021/acs.jafc.9b06909

Zhou, Z., Peng, D., Zheng, J., Guo, G., Tian, L., Yu, Z., & Sun, M. (2011). Two groups of S-layer proteins, SLP1s and SLP2s, in *Bacillus thuringiensis* co-exist in the S-layer and in parasporal inclusions. *BMB Reports*, *44*(5), 323–328. https://doi.org/10.5483/BMBRep.2011.44.5.323

Zhu, Chaohua, Guo, G., Ma, Q., Zhang, F., Ma, F., Liu, J., … Sun, M. (2017). Diversity in S-layers. *Progress in Biophysics and Molecular Biology*, *123*, 1–15. https://doi.org/10.1016/j.pbiomolbio.2016.08.002

Zhu, Chenguang, & Yu, Z. (2008). The surface layer protein of *Bacillus thuringiensis* CTC forms unique intracellular parasporal inclusion body. *Journal of Basic Microbiology*, *48*(4), 302–307. https://doi.org/10.1002/jobm.200800013

# PUBLICATIONS AND CONFERENCE PRESENTATIONS

## Publications:

1. **Hasan, M. M.,** Khan, S. N., Karim, M. M., Begum, A., & Hoq, M. M. (2019). Complete genome and plasmid sequence of a novel *Bacillus* sp. BDBD59S, a parasporal protein synthesizing bacterium. 3 Biotech, 9(9): 318.

2. Roy, S., Saha, A., Khan, S. I., **Hasan, M. M.,** Karim, M. M., Akhter, M. Z., Hoq, M. M., & Khan, S. N. (2023). Identification and Differentiation of Closely Related Members of Bacillus cereus Group by Multiplex PCR. *Bangladesh Journal of Microbiology*, *39*(1), 21–29. https://doi.org/10.3329/bjm.v39i1.64055

## Conference presentations:

1. **Hasan, M. M.,** Khan, S. N., Karim, M. M., Begum, A., & Hoq, M. M. 2019. Complete Genome and Plasmid Sequence of a Novel *Bacillus thuringiensis* strain BDBD59S, a HeLa cell killing Parasporal Crystalline Inclusion Protein Synthesizing Bacterium. 4th IPFS-GNOBB International Conference 2019, University of Dhaka, Dhaka, Nov 11-13, 2019. (Poster presentation)

2. **Hasan, M. M.,** Khan, S. N., Karim, M. M., Begum, A., & Hoq, M. M. 2019. Complete Genome and Plasmid Sequence of a Novel *Bacillus thuringiensis* strain BDBD59S, a HeLa cell killing Parasporal Crystalline Inclusion Protein Synthesizing Bacterium. The 32nd Annual Conference on Trends of Microbiology for Sustainable Agroecological Development (BSM 2019), 6th April 2019 at Jashore-7408, Bangladesh. **2nd Poster** award.

3. **Hasan, M.M.,** Saha, A., Khan, S.I., Khan, S.N., Begum, A. and Hoq, M.M. 2018. Detection of Cancer Cell Killing Activity of Parasporal Proteins of *Bacillus thuringiensis* Strains. AFOB regional symposium (the 10th ARS 2018), January 27-29, 2018, Dhaka. (Oral presentation)

4. **Hasan, M. M.,** Khan, S. N., Karim, M. M., Begum, A., & Hoq, M. M. 2018. Identification of Mosquitocidal Toxin Encoded Gene from Plasmid Sequence of indigenous *Bacillus thuringiensis* strain. The 21st National Conference and AGM, The Zoological Society of Bangladesh. December 07-08, 2018, Dhaka. (Oral presentation)

5. **Hasan, M.M.,** Khan, S.N., Begum, A. and Hoq, M.M. 2017. Studies on anticancer peptides and cytotoxicity of UMP kinase-like parasporal protein of *Bacillus thuringiensis* strain. The 3rd GNOBB-SQUARE-ACI ICBHA Conference, December 29-30, 2017, Dhaka. (Poster presentation)

6. **Hasan, M.M.,** Khan, S.N., Begum, A. and Hoq, M.M. 2017. Identification of anticancer peptides and cytotoxicity of parasporin-like transporter protein of *Bacillus thuringiensis* strains. 20th International Biennial Conference 2017, The Zoological Society of Bangladesh. December 09-10, 2017, Dhaka. (Oral presentation)

7. **Hasan, M.M.,** Khan, S.N., Begum, A. and Hoq, M.M. 2017. Finding cancer cell targeted nontoxic peptide therapeutics from the transmembrane protein of *Bacillus thuringiensis.* The 13thAsian Congress on Biotechnology (ACB 2017) "Bioinnovation and Bioeconomy", July 23-27, KhonKaen, Thailand. (Poster presentation)

8. **Hasan, M.M.,** Khan, S.N., Begum, A. and Hoq, M.M. 2017. Finding cancer cell targeted nonhemolytic peptide therapeutics from the transmembrane protein of indigenous *Bacillus thuringiensis.* 30th Annual Conference, Bangladesh Society of Microbiologists (BSM), Saturday, April 29, 2017, Dhaka, Bangladesh. (Oral presentation)

9. **Hasan, M.M.,** Khan, S.N., Begum, A. and Hoq,M.M. 2017. Identification of *ps* like Gene and *In silico* Prediction of their Anticancer Peptides from *Bacillus thuringiensis* strains of Bangladesh. 1st International Conference Global Circle for Scientific, Technological and Management Research (GCSTMR), February 4-5, 2017, Dhaka, Bangladesh. (Oral presentation)

10. **Hasan, M.M.,** Khan, S.N., Begum, A. and Hoq,M.M. 2016. Identification of *ps* genes encoding anticancer parasporin proteins from *Bacillus thuringiensis* strains of Bangladesh. National Conference 2016 on "Biochemistry and Molecular Biology for Life Sciences" Organized by: The Bangladesh Society for Biochemistry and Molecular Biology in association with Department of Biochemistry and Molecular Biology, University of Dhaka, December 10, 2016, Dhaka, Bangladesh. (Poster presentation)

# APPENDICES

## Appendix I

## Microbiological Growth Media

### Luria Bertani Agar & Broth

| Ingredients | Quantity (g/L) |
|---|---|
| Yeast extract | 5 |
| Tryptone | 10 |
| NaCl | 10 |
| Distilled water | Up to 1000 mL |
| pH | 7 |
| Agar* (for LB agar) | 15 |

**Procedure:** The ingredients are dissolved in distilled water by the process of stirring while applying mild heat. The medium is sterilized using autoclaving at a temperature of 121 °C for a duration of 15 minutes.

* In LB broth no agar was added

### Nutrient Agar

| Ingredients | Quantity (g/L) |
|---|---|
| Beef extract | 3 |
| Peptone | 5 |
| NaCl | 5 |
| Distilled water | Upto 1000 mL |
| pH | 7 |
| Agar | 15 |

**Procedure:** The ingredients are dissolved in distilled water by the process of stirring while applying mild heat. The medium is sterilized using autoclaving at a temperature of 121 °C for a duration of 15 minutes.

## Sheep Blood Agar

| Ingredient | Quantity (g/L) |
|---|---|
| Liver extract | 2.5 |
| Yeast extract | 5 |
| Proteose peptone | 15 |
| NaCl | 5 |
| Agar | 15 |
| Distilled water | Up to 1000 mL |

**Instructions:** Bring to a boil in order to fully dissolve the medium. Perform sterilization by subjecting to autoclaving at a temperature of 121 °C for a duration of 15 minutes. Cool the mixture to a temperature range of 40-50 °C and carefully add 7% v/v of sterile defibrinated blood using aseptic techniques.

## T3 Broth

| Ingredient | Quantity (g/L) |
|---|---|
| Bacto-tryptone | 3 |
| Bacto-tryptose | 2 |
| Yeast extract | 1.5 |
| $MnCl_2$ | 0.005 |
| Phosphate buffer | 50 mM |
| pH | 6.8 |
| Distilled water | Up to 1000 mL |

**Procedure:** The ingredients are dissolved in distilled water by gently swirling with the application of heat. The medium is sterilized using autoclaving at a temperature of 121°C for a duration of 15 minutes**.**

## Chromogenic agar for *Bacillus cereus* group

**Instructions:** Dissolve 49.22 g in 1 liter of distilled water. Utilize an autoclave apparatus to achieve sterilization of things by subjecting them to a temperature of 121 °C for a duration of 15 minutes, while maintaining a pressure of 15 pounds. Lower the temperature to a range of 45 to 50 °C. Aseptically introduce the rehydrated contents of 1 vial of the *Bacillus* Selective Supplement to selectively isolate *B. cereus* and *B. thuringiensis* (FD324).

# Appendix II

## Reagents and Solutions

**Phosphate buffer**

A. 0.5L of 1M $K_2HPO_4$ at 174.18 g/mol (solution A)
B. 0.5L of 1M $KH_2PO_4$ at 136.09 g/mol (solution B)

Combine the appropriate quantities of solution A and B to achieve the required pH level.

**Parasporal Protein Solubilizing Buffer (pH 11.0)**

| | |
|---|---|
| $Na_2CO_3$ | 0.53 g |
| DTT | 0.17 g |
| Sterilized $dH_2O$ | up to 100 mL |

**The Bradford Reagent**

The Bradford reagent was prepared by dissolving 100 mg of coomassie blue G-250 in 50 mL of 95% ethanol. To this solution, 100 mL of 85% (w/v) phosphoric acid was added, and the combination was further diluted to a total volume of 1 liter using water.

**Proteinase K dissolving solution (10 mL storage buffer)**

| | |
|---|---|
| Glycerol | 5 mL |
| 1M Tris-HCL | 100 µL |
| $CaCl_2$ | 0.029 g |
| Sterilized $dH_2O$ | up to 100 mL |

Add 11 mg Proteinase K within 1 mL solution and store at -20 $^oC$

**0.1 M PMSF**

| | |
|---|---|
| PMSF | 0.0174 g |
| Ethanol (95%) | 1 mL |

***Reagents for SDS-PAGE***

**30% Acrylamide-bisacrylamide Solution**

| | |
|---|---|
| Acrylamide | 29.0 g |
| Bisacrylamide | 1.0 g |
| Distilled water | 100 mL |

**Ammonium-persulphate (APS) at a concentration of 10%**

| | |
|---|---|
| APS | 1.0 g |
| Distilled water | 10 mL |
| Stored at 4 °C | |

**Buffer solution with a pH of 6.8 for use in the upper gel**

| | |
|---|---|
| Tris-base | 6.05 g |
| SDS | 0.4 g |
| pH to | 6.8 |
| Distilled water | Up to 100 mL |

**Buffer solution with a pH of 8.8 for use in the lower gel**

| | |
|---|---|
| Tris-base | 36.4 g |
| SDS | 0.8 |
| pH to | 8.8 |
| Distilled water | Up to 100 mL |

**Protein sample preparation buffer**

| | |
|---|---|
| 0.5 M tris-Cl (Upper gel buffer) | 10 mL |
| 10% SDS | 10 mL |
| 1.0 M DTT | 5 mL |
| Glycerol | 10 mL |
| Distilled water | 14 mL |

**Solution containing 0.1% Bromophenol blue (BMB) or tracking dye**

| | |
|---|---|
| $dH_2O$ | 100 mL |
| Bromophenol blue | 0.1 g |

**SDS Electrophoresis Buffer (10X)**

| | |
|---|---|
| Tris-base | 30.3 g |
| Glycine | 144.1 g |
| SDS | 10 g |
| pH | 8.3 |
| Distilled water | Up to 1000 mL |

**Staining solution (1L)**

| | |
|---|---|
| Coomassie blue | 2 g |
| Methanol | 450 mL |
| $dH_2O$ | 450 mL |
| Glacial Acetic Acid | 100 mL |

**De-staining solution (1L)**

| | |
|---|---|
| Methanol | 50 mL |
| Glacial Acetic Acid | 70 mL |
| $dH_2O$ | 880 mL |

***Phosphate-buffered saline (1X PBS)***

A solution of 1X PBS was created by dissolving 8 g of sodium chloride (NaCl), 0.2 g of potassium chloride (KCl), 1.44 g of sodium dihydrogen phosphate ($Na_2HPO_4$), and 0.24 g of potassium dihydrogen phosphate ($KH_2PO_4$) in 800 mL of distilled water (H2O). The pH was neutralized to 7.4 using hydrochloric acid (HCl) and then sterilized using autoclaving.

***Normal saline***

A solution of normal saline was created by dissolving 0.85 g of sodium chloride (NaCl) in 100 mL of distilled water and then sterilized using autoclaving.

***Bromophenol Blue (1%)***

Mix 10 mg Bromophenol Blue with 1 mL $dH_2O$ and mix well. Store at room temperature.

***Sodium dodecyl sulphate (10%)***

A solution was prepared by dissolving 10 g of Sodium dodecyl sulphate (SDS) in 80 mL of distilled water. The mixture was stirred gently on a magnetic stirrer to prevent the formation of foam. The ultimate volume was modified to 100 mL using distilled water and kept at room temperature.

***Tris-HCl solution with a concentration of 1.0 M***

A solution was prepared by dissolving 121.1 g of tris-base in 800 mL of distilled water. The appropriate pH was achieved by adding concentrated HCl, and the final volume was adjusted to 1 L using distilled water. The solution underwent sterilization using autoclaving and was thereafter kept at room temperature (RT).

***Ethylenediaminetetraacetic acid (EDTA) at a concentration of 0.5 M***

The amount of $Na_2EDTA$ (disodium ethylene diamine tetra-acetic acid) is 186.1 g and 20 g of NaOH pellets were introduced into 800 mL of distilled water and dissolved by agitation on a magnetic stirrer. The pH was modified to 8.0 by adding a small amount of 10 M NaOH and then the total volume was brought up to 1 L using distilled water. The solution was sterilized using the process of autoclaving and thereafter kept at room temperature.

***Buffer solution for Tris-Borate-EDTA (TBE) with a concentration of 10 times the standard strength.***

A solution was prepared by dissolving 108.8 g of Tris base and 55 g of boric acid in 800 mL of deionized water. Additionally, 40 mL of 0.5 M EDTA (pH 8.0) were added to the solution. The capacity was modified to 1000 mL using deionized water.

***Buffer solution with a pH of 8.0, often known as TE buffer.***

A solution of 10 millimolar tris-Cl (pH 8.0) and 1 millimolar EDTA was made by diluting concentrated stocks of 1 molar tris-Cl (pH 8.0) and 0.5 molar EDTA. The buffer was held at a temperature of 4 ºC

***Gram's Iodine Solution***

The ingredients, consisting of 1.0 g of iodine, 2.0 g of potassium iodide, and 300 mL of distilled water, are ground in a mortar and dissolved by gradually adding water. The prepared solution is thoroughly combined by agitating.

**KOH solution (16%)**

| | |
|---|---|
| KOH | 16 g |
| Distilled water | 100 mL |

**Naphthol solution (5%)**

| | |
|---|---|
| Naphthol | 5 g |
| Ethanol | 100 mL |

# Appendix III

## List of Chemicals and Reagents

| Name of chemicals/reagents/substrates | Brand/Country of Origin |
|---|---|
| Acrylamide | Carl Roth, Germany |
| Agar | Sigma, USA |
| Agarose | Promega, USA; Carl Roth, Germany |
| Aluminium sulfate | Merck, India |
| Ammonium persulphate | Wako, USA |
| Ammonium sulfate | Merck, India |
| Bis-acrylamide | Carl Roth, Germany |
| Boric acid | Merck, India |
| Bromophenol Blue | Wako, USA |
| Calcium chloride ($CaCl_2$) | Sigma, USA |
| Coomassie Brilliant Blue G250 | Thermo Scientific, USA |
| Di potassium hydrogen phosphate | Merck, Germany |
| Di sodium hydrogen phosphate | Merck, Germany |
| Dithiothreitol (DTT) | American Bio analytical, USA |
| EDTA | BDH, England |
| Ethanol (EtOH) | Merck, Germany |
| Ethidium bromide (EtBr) | Sigma, USA |
| Glacial acetic acid | Merck, Germany |
| Glycerol | Sigma, USA |
| Glycine | Wako, USA |
| HiCrome™ Bacillus Agar Base | HiMedia, India |
| Hydrochloric acid (HCl) | Merck, Germany |
| Immersion oil | Merck, Germany |
| Potassium chloride | Sigma, USA |
| Potassium dihydrogen orthophosphate ($KH_2PO_4$) | Merck, Germany |
| Lysozyme | Wako, USA |
| Methanol | Sigma, USA |
| Magnesium chloride | Sigma, USA |

| | |
|---|---|
| Magnesium sulfate | Sigma-Aldrich |
| Manganese chloride | Merck, Germany |
| n-Hexane | Merck, Germany |
| Sodium chloride (NaCl) | Sigma, USA |
| Sodium carbonate ($NaCO_3$) | Sigma, USA |
| Sodium bicarbonate ($NaHCO_3$) | Merck, Germany |
| Disodium phosphate ($Na_2HPO_4$) | Merck, Germany |
| Monosodium phosphate ($NaH_2PO_4$) | Merck, Germany |
| Sodium Hydroxide (NaOH) | Merck, Germany |
| Sodium sulfide | Sigma-Aldrich |
| Peptone | Oxoid, England |
| Phosphoric acid | Merck, Germany |
| Phenol red | Sigma, USA |
| PMSF (phenylmethylsulfonyl fluoride) | Thermo Scientific™ |
| Protease Inhibitor | |
| Potassium hydrogen phosphate ($K_2HPO_4$) | Sigma, USA |
| Protein marker | BenchMark™ Protein Ladder (Life Technologies), Precision plus protein standards (All blue), Bio-Rad, USA; Pre-stained protein Marker, NEB, England |
| Proteinase K | MP Biomedicals LLC, France |
| Sodium acetate | Merck, Germany |
| Sodium dodecyl sulphate (SDS) | Wako, USA |
| Sodium hydroxide (NaOH) | Sigma, USA |
| Gibco Trypsin-EDTA (0.25%), phenol red | Thermo Fisher Scientific |
| Tris-base | Sigma, USA |
| Tryptone | BD, USA |
| Tryptose | BD, USA |

# Appendix IV

## List of Equipment's

| Equipment's | Model/ Brand/Country of Origin |
|---|---|
| Autoclave machine | Hirayama model HL-42, AE, Japan |
| Biosafety cabinet | ESCO Class II BSC, USA |
| Centrifuge machine | MX-305 High Speed Refrigerated Micro Centrifuge |
| DNA Ladder | GeneRuler 100 bp DNA Ladder (Invitrogen, USA) |
| Electronic balance | SHIMADZU, ELB200 |
| Glassware sterilizer | Redline, Binder, Germany |
| Microbiological incubator | Redline, Binder, Germany |
| Magnetic stirrer | CIMAREC, Barnsted Thermolyne |
| Micropipettes | Eppendorf research and Nichiryo |
| Nanodrop spectrophotometer | Nanodrop 2000, Thermoscientific; NanoDrop One, Thermoscientific |
| Sonication machine | Omni-Ruptor 4000 Ultrasonic Homogenizer, OMNI International. USA |
| Orbital shaker incubator | Excella E25 Incubator Shaker series, New Burnswick Scientific |
| pH meter | INOLAB WTW series, pH 720 |
| Power supply | BIORAD, USA |
| Refrigerator (4 $^{o}$C) | Royal Frestech |
| Spectrophotometer | Genesys 5, Thermospectronic |
| Thermal cycler | ProFlex, Applied Biosystems |
| Thermo stated shaking water bath | Water Bath 1083, GFL |
| Vortex mixture | VM-2000, DIGISYSTEM |