ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

BIOLOGIA CELLULARE E MOLECOLARE

Ciclo 38

Settore Concorsuale: 05/I1 - GENETICA

Settore Scientifico Disciplinare: BIO/18 - GENETICA

FUNCTIONAL AND STRUCTURAL ANALYSIS OF TnBVANK1, AN IκB-LIKE PROTEIN

Presentata da: Mattia Longanesi

Coordinatore Dottorato

Anna Maria Porcelli

Supervisore

Giovanni Perini

Co-supervisore

Roberto Bernardoni

Esame finale anno 2026

CONTENTS

1. INTRODUCTION……...3

1.1. Importance of crop protection…….....3

1.2. Mechanism of action of conventional insecticides…….....5

1.3. Problems related to the use of neuroactive insecticides.....6

1.4. Need for sustainable alternatives.....7

AIM OF THE THESIS.....9

1.5. Parasitoid/host system *Toxoneuron nigriceps-Heliothis virescens*.....10

1.6. *Tn*BV genome and ANK proteins.....11

1.7. *Tn*BVANK1 protein.....14

1.8. NF-kB immune system pathway.....15

1.9. *Drosophila melanogaster* as a model system.....17

1.10. *Tn*BVANK1 expression in *Drosophila melanogaster*.....17

2. MATERIAL AND METHODS.....18

2.1. Fly strains.....18

2.2. TARGET system.....19

2.3. Survival test.....20

2.4. RNA extraction from adult flies.....21

2.5. Reverse transcription.....22

2.6. qPCR analysis.....23

2.7. Expression of *Tn*BVANK1.....25

2.8. Purification of *Tn*BVANK1.....27

2.9. SDS-PAGE assay.....29

2.10. CD-spectroscopy analysis.....30

2.11. NMR analysis.....30

3. RESULTS.....31
3.1. The effect of *Tn*BVANK1 ubiquitous expression on the survival of adult flies.....31
3.2. Ubiquitous expression of *Tn*BVANK1 in adult *Drosophila melanogaster* induces an increased transcript levels of five antimicrobial peptides genes.....32
3.3. *Tn*BVANK1 was correctly express in *E.coli* BL21.....35
3.4. Purification of *Tn*BVANK1 by affinity chromatography (IMAC) and by size exclusion chromatography (SEC).....36
3.5. Evaluation of *Tn*BVANK1 expression by SDS-PAGE and mass spectrometry analysis.....37
3.6. Analysis of the structure of *Tn*BVANK1 by circular dichroism spectrometry (CD-spectroscopy) and nuclear magnetic resonance spectrometry (NMR).....39
4. DISCUSSION.....43
5. REFERENCES.....45

1. INTRODUCTION

1.1 importance of crop protection

The protection of the crops and of their products against insects has old roots, dating back to the origin of agriculture. Until the mid-1800s, insect control depended largely on picking or washing off the pests. The few chemical weapons included inorganic sulfur (since 1000 bc), arsenic (900 ad), and later lead arsenate, cryolite, and boric acid, some of which (e.g. sulfur) are still nowadays major pest control agents. These inorganics compounds are general poisons, used in large amounts, often for marginal control (Casida et al., 1998).

Subsequently, botanical insecticides were discovered which attracted greater interest from researchers due to their greater selectivity and efficiency. These insecticides include pyrethrum, rotenone, ryanodine, veratridine, and azadirachtin, but they were limited in availability and generally too expensive and photolabile.

In the 1930s, following screening of industrial compounds of unknown biological activity against specific pest species, several synthetic organic compounds with insecticidal activity were discovered. In 1939, Paul Hermann Müller condensed chlorobenzene and chloral in the same way that, another researcher, Othmar Zeidler reported in 1874 for the preparation of DDT. Müller was searching for an insecticide and submitted his compounds to the screening program of Geigy AG in Basel, Switzerland. The results showed that DDT was active on many pests and controlled them for long periods of time. DDT proved to be so effective that the spread of malaria was also drastically reduced (Casida et al., 1998) (Fig. 1).

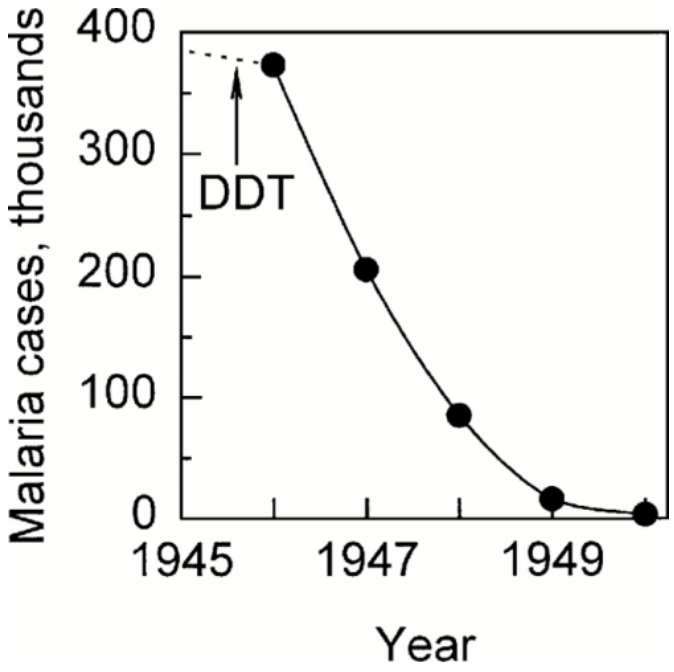

Figure 1: Trend of malaria spread compared to DDT diffusion.

Due to the high efficiency of DDT, other chlorinated compounds were examined with the goal of discovering new, cheap, and persistent contact insecticides. Hexachlorocyclohexane (first synthesized by Faraday in 1825) and its isomer, lindane, were found in 1942 to be effective insecticides. The median entry year for the leading chlorinated products was 1947, and during a 40-year period more than 8,000,000,000 pounds were used.

Starting in the early 1970s, these compounds were severely restricted or banned in the United States, with varying levels of curtailment in other parts of the world due to toxicological problems and environmental contamination. However, lindane is still an important compound, and even DDT, is a major product in certain parts of the world. Thus, the chlorinateds at one time dominated and still play a major role in pest insect control.

For the reasons listed above, synthetic pyrethroids began to spread as a new class of insecticides. These insecticides were synthesized from pyrethrins and achieved remarkable efficacy, setting new standards for contact insecticides (Casida et al., 1998).

Research on synthetic pyrethroids yielded the following results

1. Photostability without compromising biodegradability.
2. Selective toxicity conferred by target-site specificity.
3. Modification of every part of the molecule with retention of activity.
4. Maintenance of high insecticidal potency while minimizing fish toxicity.
5. Development of compounds effective as fumigants and soil insecticides.
6. Optimization of potency to allow corresponding reduction in environmental contamination.

Chlorinated insecticides such as DDT and synthetic pyrethroid insecticides act through the same molecular mechanisms, which we will discuss in the next paragraph.

1.2 Mechanism of action of conventional insecticides

To understand the mechanism of action of these classes of insecticides, we must start with a fundamental concept: The action potential.

The action potential is the rapid short-lasting change in the electrical potential across the membrane of a cell, primarily neurons and muscle cells, that allows for the transmission of signals. It's a fundamental process for communication within the nervous system and for initiating muscle contractions.

A neuronal action potential has three main stages: depolarization, repolarization, and hyperpolarization. The initial depolarization is determined by the cell's threshold voltage, the membrane potential at which voltage-gated sodium channels (Na^+) open to allow an influx of sodium ions. The flow of positive sodium ions into

the cell leads to further depolarization of the membrane, thus opening more Na^+ in a positive-feedback loop. Depolarization in mature neurons lasts approximately 1 msec, at which time the Na^+ are inactivated and no longer able to flux ions.

Repolarization begins when voltage-gated potassium channels (K^+) open. Although K^+ have approximately the same threshold voltage as Na^+ the kinetics of the potassium channel are much slower. Therefore, after approximately 1 msec, there is an opening of the slower K^+ channels that is coincident with the inactivation of the faster Na^+ channels. The flow of potassium ions out of the cell results in a decrease in membrane potential towards the cell's resting voltage. When the membrane potential falls below the threshold, both the Na^+ and the K^+ begin to close. However, the K^+ have slow kinetics and remain open slightly longer to return the cell to resting membrane voltage. The brief dip in the membrane potential below the normal resting voltage is termed hyperpolarization. To reestablish the appropriate balance of ions and the normal membrane potential, an ATP-driven pump (Na/K-ATPase) induces movement of sodium ions out of the cell and potassium ions into the cell. The three events of the action potential are best described in the following graph (Fig. 2).

.

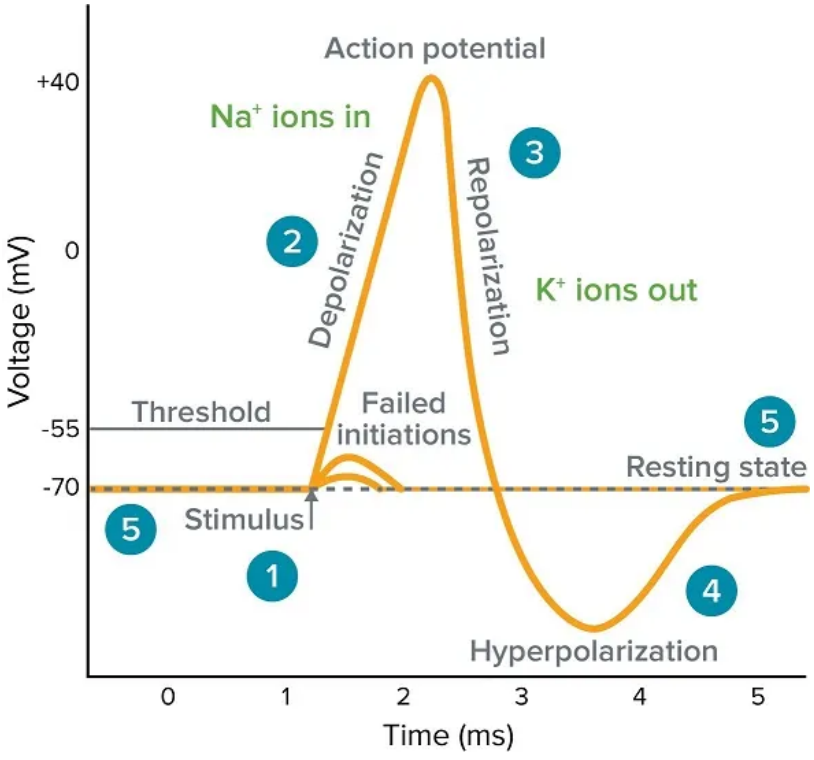

Figure 2: Overview of the action potential graph.

The DDT – type insecticides act primarily on the peripheral nervous system. At the sodium gates of the axon, DDT exerts its toxic action by preventing the deactivation or closing of that gate after their activation and

membrane depolarization. The result is a lingering leakage of Na^+ ions through the nerve membrane, creating a destabilizing positive potential.

The hyperexcitability of the nerve results in trains of repetitive discharges in the neuron after a single stimulus and these discharges occur spontaneously. This hyperexcitability induces tremors and ultimately, paralysis and death in insects.

1.3 Problems related to the use of neuroactive insecticides

The early neuroactive insecticides like DDT and syntethic phyretroids were effective, inexpensive, and often persistent, ideal properties for the selection of resistant strains. Resistance can be conferred by a minimal (or even single) amino acid substitution at the target site. (Casida et al., 1998). Resistance problems with this first generation of neuroactive insecticides prompted the search for new neurotoxicants, like imidacloprid, acting at different sites to circumvent cross resistance. This compound is a neonicotinoid insecticide and acts as an agonist of the post – sinaptic acetylcholine receptor (nAChR). Acetylcholine is a key neurotransmitter for communication between neurons and between neurons and muscles. It controls voluntary muscles, intestinal peristalsis, glandular secretion, and cardiac function. It is also essential for memory, attention, and learning in the central nervous system. It is synthesized from choline and acetyl - CoA and is degraded by the enzyme acetylcholinesterase. Acetylcholine binds to cholinergic receptors (nicotinic and muscarinic) present on the postsynaptic cell to depolarize it and transmit the nerve impulse. Finally, after carrying out its function, acetylcholine is cleaved in choline + acetate by the enzyme acetylcholinesterase (AChE) (Casida et al., 1998). Neonicotinoid insecticides mimics acetylcholine, but unlike it, they are not broken down by acetylcholinesterase. This persistent binding to acetylcholine receptors (nAChRs) on the postsynaptic neuron leads to overstimulation of the impulse resulting in prolonged nerve activation, paralysis, and eventual death of the insect. No target site – associated cross - resistance problems exist and, so, the synthetic nicotinoids are outstanding replacements the first neuroactive insecticides in many uses for sucking insect pests of crops (Casida et al., 1998).

1.4 Need for sustainable alternatives

Insecticides have been, and still are, an important tool in intensive agriculture, but it was discovered that they often caused fish kills or bird mortality. Furthermore, the release of chemical pesticides into the environment causes contamination of the surrounding environment and the soil. The active ingredients of the chemical compounds, once released into the soil, can penetrate deep into the groundwater, contaminating groundwater. This is a major problem because, once in groundwater, the chemicals can enter the public water system and ultimately end up on consumers' tables. However, in 1940s and 1950s these effects were considered as an unavoidable side effects associated with pest control. Over the years, this mortality rate became unacceptable, highlighting all the limitations of the approach using chemical insecticides.

Pest control strategies and procedures, like many aspects of biology and agriculture, have changed significantly thanks to advances in molecular biology and genetic engineering. These new methodology has offered for the first time the opportunity to produce bioinsecticides without chemical compound. Even more importantly, genetic engineering made it possible to produce insecticidal proteins and peptides and deliver them to insects via bacterial, baculoviral, or plant - based systems. Insecticide candidates were no longer limited to low-molecular-weight, fat-soluble chemicals, but could now include, for the first time, large, water-soluble molecules. The challenge became to select the best toxins, enzymes, or hormones for pest control and then develop practical production and delivery systems.

An example of a toxin is the toxin produced by the bacterium *Bacillus thuringesis*. This bacterium is a member of the *Bacillus cereus* group that also includes *B. cereus*, *B. anthra-cis*, and *B. mycoides* (Helgason et al., 2000). The feature that distinguishes *B. thuringiensis* from the other members of the *B. cereus* group is its entomopathogenic properties. This bacterial species produces insecticidal proteins (d-endotoxins) during sporulation phase as parasporal inclusions, which predominantly comprise one or more proteins, called Cry and Cyt toxins. These protein toxins are highly selective to their target insect, are innocuous to humans, vertebrates, and plants, and are completely biodegradable. Therefore, *B.thuringiensis* is a viable alternative for the control of insect pests in agriculture.

Numerous *B. thuringiensis* strains have been isolated that show activity towards lepidopteran, dipteran, or coleopteran insects (Schnepf et al., 1998). In recent years, *B. thuringiensis* strains active against Hymenoptera, Homoptera, Orthoptera, and Mallo-phaga insect orders and to other non-insect organisms like nematodes, mites, and protozoa have been isolated (Crickmore et al., 1998; Wei et al., 2003). The entomopathogenic activity of *B. thuringiensis* is mainly due to Cry toxins. One feature that distinguishes these Cry proteins is their high selectivity for the target insect. More than 200 different *cry* genes have been isolated, and this constitutes an import arsenal for the control of a wide variety of insect pests. In this chapter we will summarize the present knowledge of the pathogenic properties of *B. thuringiensis,* the mode of action of Cry toxins, and their application in agriculture and disease-vector control.

The primary action of Cry toxins is to lyse midgut epithelial cells in the target insect by forming lytic pores in the apical microvilli membrane of the cells (Gill et al., 1992; Schnepf al., 1998; Aronson and Shai, 2001; de Maagd et al., 2001). In order to exert its toxic effect, Cry pro-toxins in the crystalline inclusions are modified to membrane-inserted oligomers that cause ion leakage and cell lysis. The crystal inclusions ingested by susceptible larvae dissolve in the alkalyne pH of the environment of the gut, and the solubilized inactive pro-toxins are cleaved by midgut proteases yielding 60–70 kDa protease resistant proteins (Choma et al., 1990). Toxin modification results in their activation and involves the proteolytic remove of an N-terminal peptide. The activated toxin then binds to specific receptors (APN receptors) on the brush border membrane of the midgut epithelium columnar cells of susceptible insects (Schnepf et al., 1998; de Maagd et al., 2001) before inserting into the membrane (Schnepf et al., 1998; Aronson and Shai,2001). Toxin insertion leads to the formation of lytic pores in microvilli apical membranes (Schnepf et al., 1998; Aronson and Shai,

2001). Cell lysis and disruption of the midgut epithelium releases the cell contents providing to the spores a rich medium that is suitable for spore germination, and leading to a severe septicemia and insect death (Schnepf et al., 1998; de Maagd et al., 2001).

Using genetic engineering and molecular biology techniques, expression vector have been created in which the Cry gene sequence of *Bacillus thuringensis* has been cloned into the multi cloning site (MCS) of these plasmid. This vector are transformed into crops to produce insect-resistant crops, such as the *Bacillus thuringensis* corn (Bt corn). This corn is called Bt corn because it has been transformed with an expression vector that expresses the Bt toxin.

Bt toxin represents a valid alternative to chemical insecticides, however, it has been observed that its constant use has selected insects resistant to the toxin. One of the major concerns regarding the use of *B. thuringiensis* is the generation of insect populations resistant to Cry toxins. Resistance to these toxins could be obtained by interfering with any of the steps involved in the mode of action of Cry toxins. In particular, the most frequent mechanism of resistance to Cry toxins is due to the altered receptor binding (Ferre´and VanRie, 2002).

The specific goal of insecticide research is to discover, develop, and understand new products and methods for safe and effective pest control, maximizing food production and public health. In order to avoid the selection of resistant insects, the new research has also boosted to the identification of new control strategies based on the use of natural antagonists, which has favored the successful consolidation of classical biological control. In particular, the studies on the antagonistic associations between insects and their natural enemies, has generated some information on the molecular interactions, which offers the opportunity to develop “bio-inspired” pest control strategies.

AIM OF THE THESIS

Plant diseases cause considerable losses in crop production and storage. Nowadays, growers still rely heavily on chemical pesticides to prevent, or control these diseases. However, the high effectiveness and ease of utilization of these chemicals can result in environmental contamination and the presence of pesticide residues on food, in addition to social and economic problems. Consequently, there is an increasing demand from consumers and officials to reduce the use of chemical pesticides. In this context, biological control through the use of natural antagonistic organisms has emerged as a promising alternative. Biopesticide is a mass-produced, biologically based agent manufactured from a living organism or a natural product and which is sold for the control of plant pests. The agents used as biopesticides are usually broken down into two categories:

1. Natural organisms (bacteria, parasitoid, fungi, oomycetes, and viruses).
2. Biochemical extracted from plant such as essential oils, and various compounds synthesized by other organisms such as chitin and chitosan.

These biopesticides present many advantages in term of sustainability, mode of action and toxicity compared to chemical pesticides. In particular, the impressive diversity of Hymenoptera, which represent the largest group of parasitoid insect antagonists, include nearly 10–20% of all insects and is the largest reservoir of molecular biodiversity, that can offer a wide selection of virulence factors, having potential insecticide activity against a number of insect species. The host-parasitoid associations in basal evolutionary lineages of Hymenoptera are characterized by the presence of venom blends which contain active components causing rapid and irreversible paralysis, used by idiobionts to block the development of their hosts, while more derived lineages, often showing endophagy, are koinobionts that regulate a number of physiological traits of their hosts, which continue to grow and are not paralyzed/suppressed. Among these latter, there is a group of ichneumonoids wasp endophagous parasitoids of lepidopteran larvae, which harbor a symbiotic virus of the *Polydnaviridae* family.

In order to create a suitable environment for the development of the parasitoid's progeny, Polydnaviruses (PDV) are among the most potent viral immunosuppressors existing in nature, encoding host virulence factors, which are able to modulate host physiology, by altering its vital functions, in order to create a suitable environment for the development of the parasitoid's progeny. Virulence factors include several proteins such as the ANK proteins.

The aim of my PhD project is to study one of the virulence factors of the parasitoid ichneumonid wasp *Toxoneuron nigriceps*, which parasitizes its natural host *Heliothis virescens.* In particular, the project concerns the functional analysis of an ANK protein, *Tn*BVANK1 in *Drosophila melanogaster* with the aim of evaluating its possible use as a new generation of eco-sustainable bioinsecticide.

1.5 Parasitoid/host system *Toxoneuron nigriceps-Heliothis virescens*

Parasitoids insects are divided in several orders of insects (Diptera, Coleoptera, Lepidoptera, Trichoptera, Neuroptera, Strepsiptera), but they are especially common in the Hymenoptera, for which recent estimates suggest that 10% to 20% of the Hymenoptera insects may be parasitoid wasps. Furthermore, most parasitoids are restricted to attacking a particular host life stage (eggs, larvae, pupae, or adults).

In particular, *Toxoneuron nigriceps* (Hymenoptera, Braconidae) is an endophagous parasitoid of the larval stage of the tobacco budworm *Heliothis virescens* (Lepidoptera, Noctuidae). Parasitized *H. virescens* larvae are developmentally arrested and show a complex array of pathological symptoms ranging from the suppression of the immune response to a disruption of the host endocrine system. (Fig. 3) Most of these pathological syndromes are induced by the polydnavirus associated with *T. nigriceps* (*Tn*BV) that actively replicating in the calyx region of the wasp's ovary (Pennacchio et al., 2006, Varricchio et al., 1999).

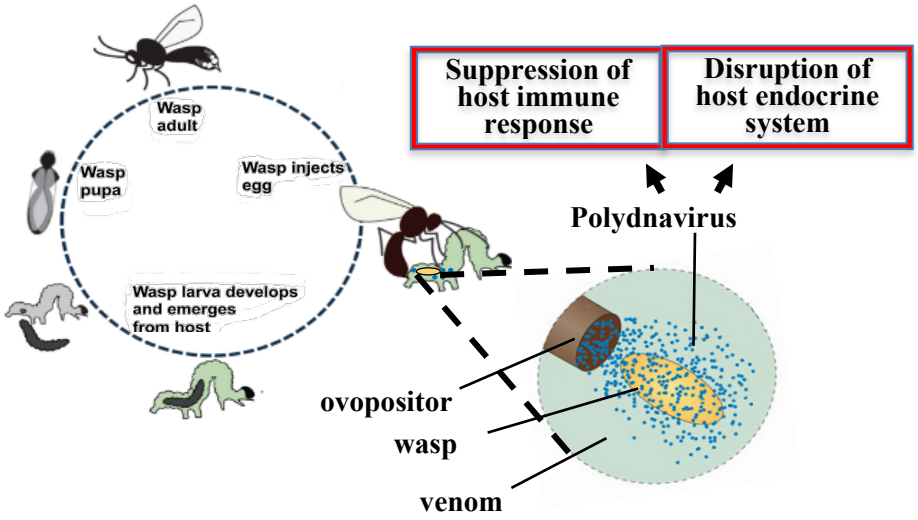

Figure 3: Toxoneuron nigriceps parasitization mechanism and its effect on the Heliothis virescens larva.

During oviposition, *T. nigriceps* females inject the egg into the host larva, along with the venom and the ovarian calyx fluid that include the symbiotic polydnavirus (PDV) named T. *nigriceps* bracovirus (*Tn*BV) (Fig. 4A), which contain the segmented viral genome (Fig. 4B). The parasitoid embryo (Fig. 4C) is surrounded by a serosal membrane that, at hatching, dissociates into single rounded cells, the teratocytes (Fig. 4D). These cells freely float in the host haemocoel, and, as the parasitoid larva develops, they increase in size and ploidy level, without undergoing division (Dahlman and Vinson, 1993, Pennacchio et al., 1994). *Tn*BV and teratocytes are the two major host regulation factors, respectively of maternal and embryonic origin, responsible for the physiological alterations induced by parasitization (Tanaka and Vinson, 1991, Pennacchio et al., 1998).

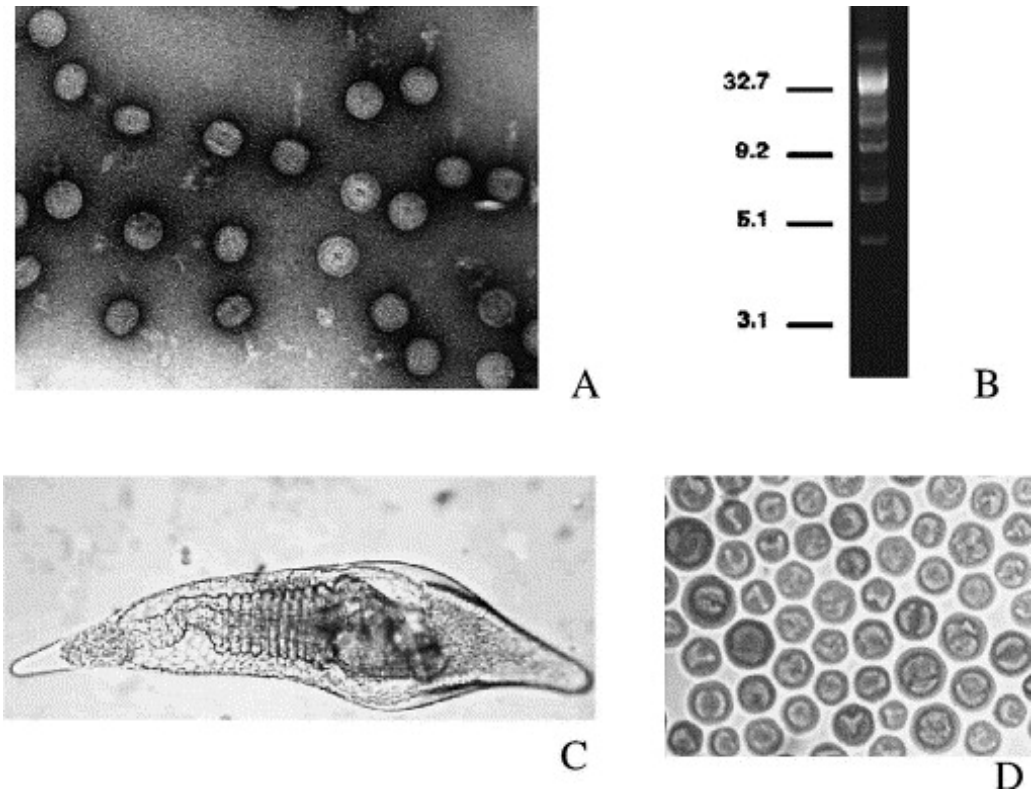

Figure 4: (A) Negative staining of liberated nucleocapsid of Toxoneuron nigriceps polydnavirus (TnBV). (B) Undigested TnBV DNA extracted from purified viral particles, stained with ethidium bromide and separated by 0.7% agarose gel electrophoresis. (C) T. nigriceps embryo; at the hatching, the terminal poles of the embryonic serosa dissociate into single cells called teratocytes (D).

1.6 *Tn*BV genome and ANK proteins

The *TnBV* genome is composed of DNA circles of different size, ranging from 3.4 to 13.3 kb. The coding sequences include open reading frames (ORFs) encoding putative proteins that for the majority belongs to 4 gene families: ANK protein with ankyrin repeat motif (Falabella et al., 2007), PTP (protein tyrosine phosphatase) (Provost et al., 2004; Falabella et al., 2006), UDP (glucose-6-dehydrogenaseand sugar transporter) and BEN domain proteins (Falabella et al., 2006).

In synergy with other parasitic factors, the symbiotic PDV seems to play the major role in inducing the alteration of host physiology. In particular, the injection and active transcription of each *Tn*BV genes family are responsible for the alteration of the natural homeostasis of the host as explained in (Fig. 3).

In my PhD research, I dealt with *Tn*BVANK1, one of the ANK proteins expressed by the virions of the polydnavirus TnBV. Three ankyrin-like open reading frames (ORFs) were identified in the *Tn*BV genome, and they encode for *Tn*BVANK1, *Tn*BVANK2 and *Tn*BVANK3 (Falabella et al., 2007) (Fig 5). All *Tn*BVANK proteins contain four ankyrin domain, but the phosphorylable serine residue in the SRD domain, essential for the IKB protein function, are absent (Falabella et al., 2007). Moreover, all of putative proteins encoded by the *TnBVank* genes lack the C-terminal PEST domain, characteristic in all active IκB like proteins, which is essential for protein degradation (Ghosh et al., 1998, Ghosh and Karin, 2002; Webb and Strand, 2006) and for the protein turnover control (Rogers et al., 1986).

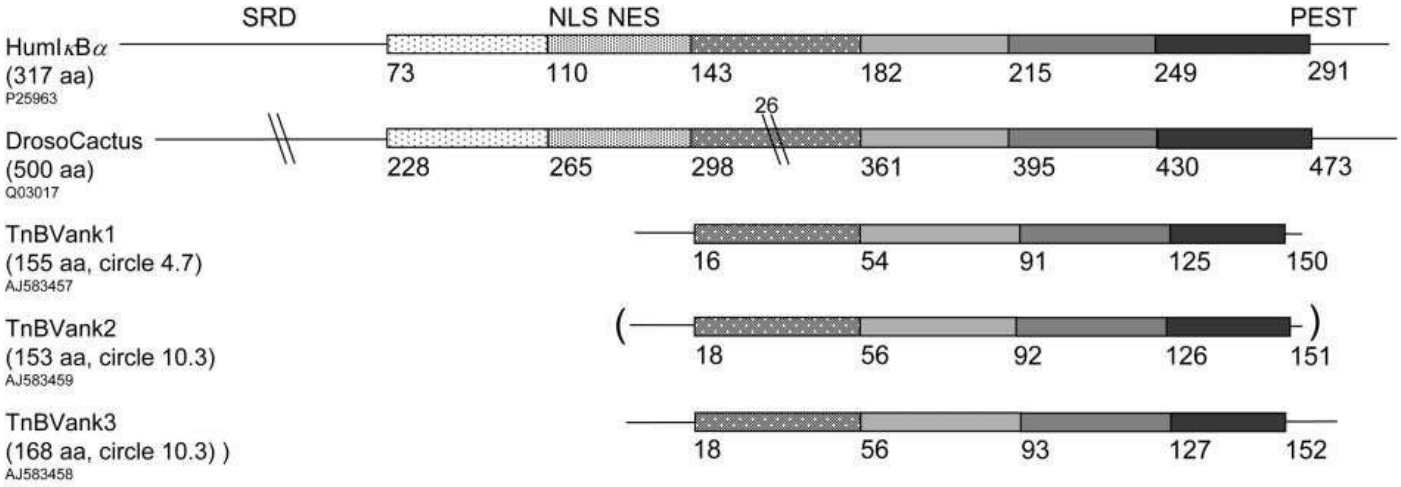

Figure 5: Schematic representation of the proteins encoded by TnBV ank compared with the ankyrin repeats of human (Hum) IkBa and Drosophila melanogaster (Droso) Cactus (GenBank accession numbers are indicated below the name ofthe proteins). For the latter, an insertion of 26 aa in the ank3 repeat, conserved across insects, is denoted with a double slash. The numbers under each protein scheme indicate the amino acid positions delimiting the different ankyrin repeats.

Ankyrin repeat proteins are found in all three kingdoms of life. Fundamentally, these proteins are involved in protein-protein interaction in order to activate or suppress biological processes. The basic architecture of

these proteins comprises repeating modules forming elongated structures. Due to the lack of long-range interactions, a graded stability among the repeats is the generic properties of this protein family determining both protein folding and biological function (Kumar A, Balbach J. Folding and Stability of Ankyrin Repeats Control Biological Protein Function. Biomolecules. 2021).

Among the repeat protein families, the ANK repeat proteins form the largest family, abundantly found in bacteria, archaea, eukaryote, and viral genomes. Ankyrin repeats are part of diverse proteins that share the common function to mediate protein - protein interactions, for example as membrane bound proteins (e.g., notch membrane receptor), as cytosolic proteins, or proteins in the nucleus (inhibitor subunit (IkBα) of NFkB transcriptional factor) (Fig. 6).

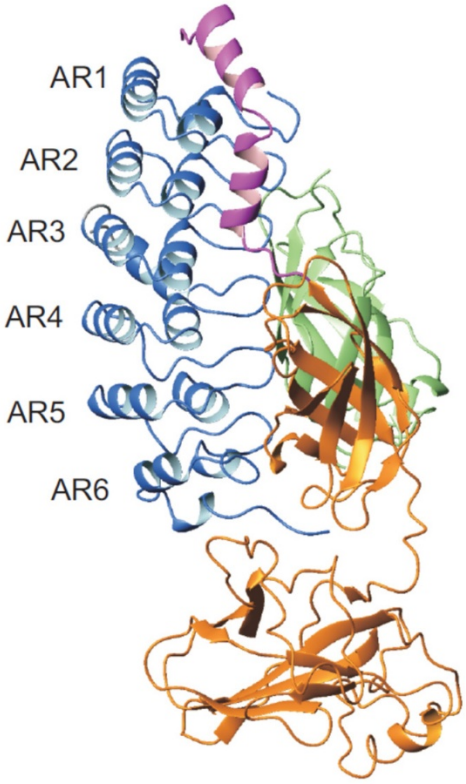

Figure 6: Structure representation of the IκBα ankyrin domain in complex with NF-κB. The IκBα ankyrin domain, composed of six repeats, is illustrated in blue, and the two monomers of NF-κB p50 in orange (at the C-terminal) and the p65 in green (at the N-terminal). Furthermore, NLS sequence of p65 is illustrated in magenta (PDB ID: 1NFI). Specifically, the p65 monomer contains the DNA-binding domain and a dimerization domain (RHR) that also contains the nuclear localization signal (NLS) of NF-kB. The ankyrin domains of Iκbα have the ability to interact with the RHR domain of p65, 'masking' the NLS signal and preventing the translocation of NF-kB into the nucleus.

Most proteins contain tandem arrays of helices in 2–7 repeats and up to 33 repeats were found in a single protein. The consensus sequence of an ankyrin repeat comprises 33 residues, and deviations from this motif typically reduce the thermodynamic stability of the respective repeat. Crystal and NMR structures of the ankyrin repeats reveal a highly conserved structural motif. The ankyrin repeat is characterized by a pair of antiparallel α-helices that is linked to its neighboring ankyrin repeat via an anti-parallel β-loop, with the first

β-strand of the β-loop contributed by the C-terminus of the repeat at position i and the second β-strand contributed by the N-terminus of the ankyrin repeat at position i+l. Tandem ankyrin repeats stack approximately in parallel, so that the helices in one repeat pack against their counterparts in adjacent repeats. The β-strands are roughly perpendicular to the axis of the helices, giving the stack an L - shaped cross-section. There is a left-handed twist to ankyrin repeats and the tandem repeat stack is slightly curved, creating concave and convex faces. Often, the β-loop region can form a continuous β-sheet. One single ankyrin repeat is unable to adopt a folded structure due to its intrinsic instability. Therefore, at least 2 repeat are required as minimum folding units for the formation of a hydrophobic core to overcome chain entropy. The mode of protein-protein interactions that are mediated by ankyrin repeats is highly conserved. In all complexes, contacts between the ankyrin repeats and target proteins involve the β-loop fingers of the ankyrin repeats (Kumar A, Balbach J. Folding and Stability of Ankyrin Repeats Control Biological Protein Function. Biomolecules. 2021).

These molecular features of the viral ankyrin gene products suggest strongly that they may disrupt cell pathways by binding host NF-kB molecules irreversibly. Moreover, it has been shown that NF-kB/Rel immunoreactive proteins are actually sequestered in the cytoplasm of the cells of parasitized larvae (Falabella et al., 2007). ANK proteins are most probably involved in this sequestration, in which they may be aided by other factors such as ovarian proteins, venom components and other PDV products. In parasitized lepidopteran larvae, which are unable to encapsulate foreign invaders, the PDV-induced disruption of NF-kB signalling probably affects both the humoral and cellular immune responses. The negative impact of NF-kB signalling disruption on antimicrobial response, which has been reported in several lepidopteran host–parasitoid associations (Shelby & Webb, 1999; Shelby et al., 1998), could be interpreted as an efficient suppressive strategy of the host cellular immune response.

1.7 *Tn*BVANK1 protein

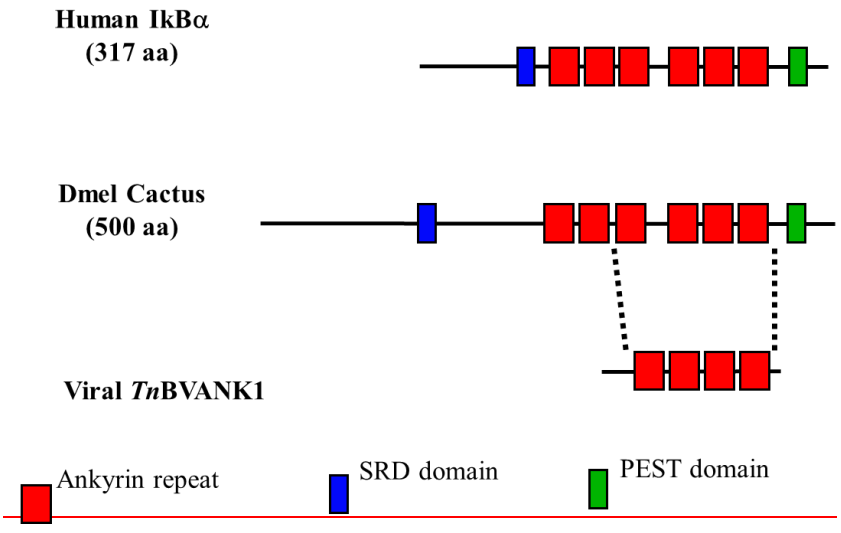

Figure 7: Overview of the protein domains of Iκbα, Cactus, and TnBVANK1.

A research published in 2014 showed that transgenic tobacco plants expressing *Tn*BVANK1 showed insecticide activity and caused developmental delay in *Spodoptera littoralis* larvae feeding on them. This effect was more evident in a transgenic line showing a higher number of transcripts of the viral gene (Di Lelio et al., 2014). Immunolocalization experiments evidenced the accumulation of this viral protein in the midgut, where it formed a thick layer coating the brush border of epithelial cells. In vitro transport experiments demonstrated that the presence of recombinant *Tn*BVANK1 exerted a dose-dependent negative impact on amino acid transport. These results open new perspectives for insect control and stimulate additional research efforts to pursue the development of novel bioinsecticides, encoded by parasitoid-derived genes. However, future work will have to carefully evaluate any effect that these molecules may have on beneficial insects and on non-target organisms.

Furthermore, in parasitised *H.virescens* larvae, after bacterial challenge, the nuclear import of NF-κB was inhibited (Falabella et al., 2007). Moreover, the expression of *Tn*BVANK1 gene in human HeLa cells reduced the efficiency of the TNF-α-induced expression of a reporter gene under NF-κB transcriptional control, strongly suggesting that *Tn*BVANK1 protein is involved in the suppression of the insect immune response (Bitra et al., 2012; Falabella et al., 2007). The early expression of the *TnBVank1* gene (three hours after parasitism) in haemocytes of *H. virescens* last instar larvae and the presence of transcripts several hours after parasitism, corroborate the *Tn*BVANK1 involvement during early immune response suppression of parasitised host, also through the activation of cell death attributable to a process of apoptosis (Falabella et al., 2007).

As reported before, (Salvia et al., 2017), *Tn*BVANK1 induces cell death in also in *Drosophila* S2 cells that stably express *Tn*BVANK1. Furthermore, coimmunoprecipitation and coimmunolocalization experiments in *Drosophila* S2 cells, showed that *Tn*BVANK1 interacts with ALG-2 interacting protein X (Alix). Alix is Ca^{2+} - binding protein necessary for cell death and it was first characterized as interactor of apoptosis. (Missotten et al., 1999). It was demonstrated that *Tn*BVANK1-Alix interaction is required to induce apoptosis in all cells expressing this gene, both in vitro (S2 cells) and in vivo (*H. virescens* larval haemocytes) (Salvia et al., 2017). Indeed, upon Alix silencing by RNA interference (RNAi), *Tn*BVANK1was no longer able to induce apoptosis (Salvia et al., 2017).

As explained before, and also as represented in (Fig. 7), *Tn*BVANK1 lacks the regulatory sites, serine residues at the N-terminus and PEST domain at the C-terminal, and only has only 4 ankyrin domains. These peculiar features and the scientific literature data that I have listed here, strongly suggest that viral ankyrins may act as inhibitors by irreversibly binding NF-kB/Rel transcription factors and retaining it in the cytoplasm, altering the signal transduction cascade, and affecting a variety of cellular responses in parasitised hosts, including alteration and the suppression of innate immune defenses. (Falabella et al., 2007)

1.8 NF–*kB* immune system pathway

To better understand the interaction between *Tn*BVANK1 and NF-κB, it is necessary to understand the NF-κB molecular pathway (Fig. 8).

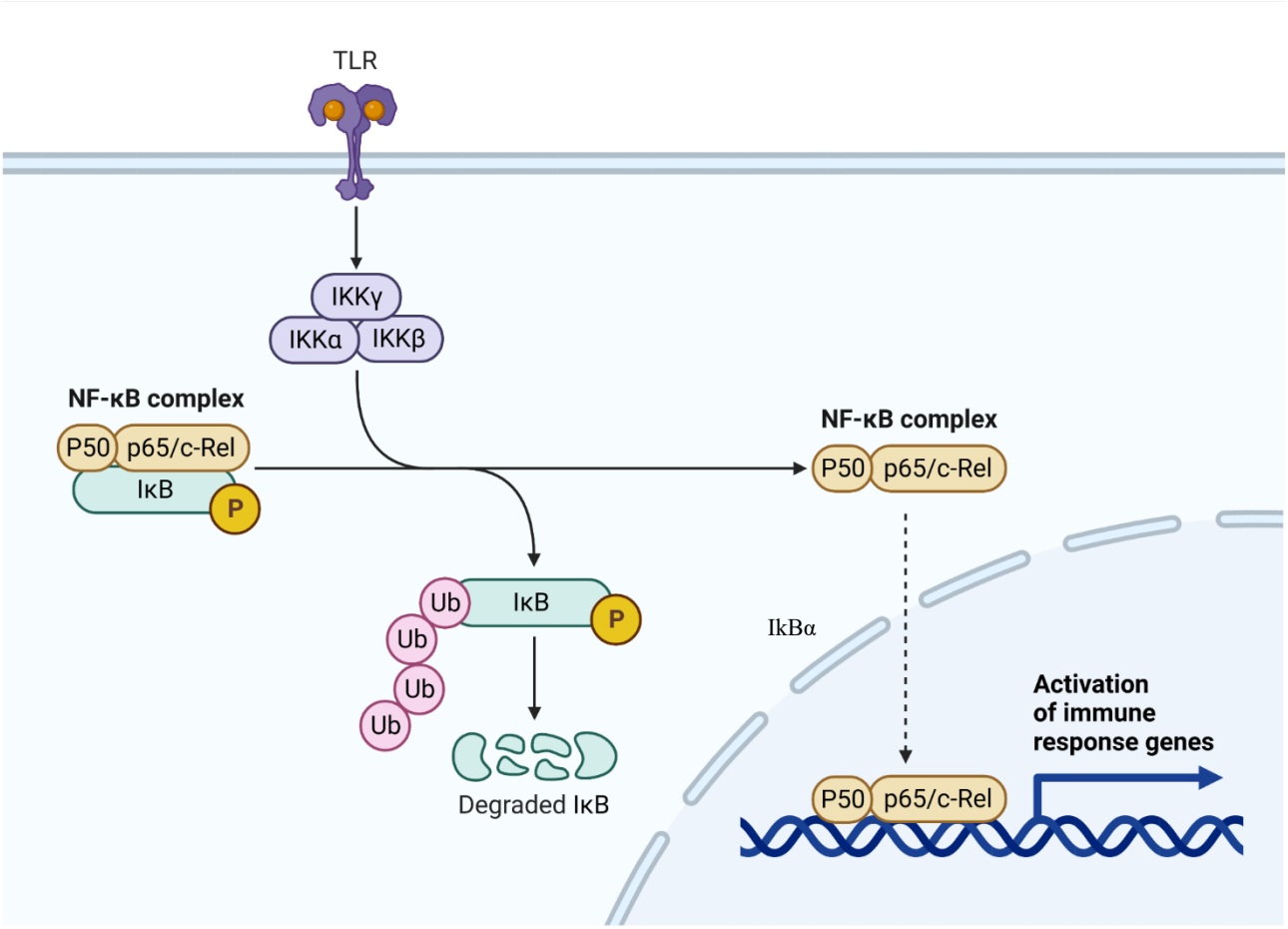

Figure 8: Overview of the NF - κB immune system pathway.

As it can be seen from (Fig. 8), in the absence of immunostimulation, the transcription factor complex NF-kB interacts with its inhibitor Ikb blocking the traslocation of NF-kB into the nucleus, the subsequent activation of the immune response and the transcription of antimicrobial peptides. The IκB proteins mask the nuclear localization signals (NLS) of the NF-κB proteins keeping them sequestered in an inactive state in the cytoplasm. Activation of the NF-κB is initiated by the signal-induced degradation of IκB proteins. This occurs primarily via activation of a kinase called the IκB kinase (IKK). IKK is composed of a heterodimer of the catalytic IKKα and IKKβ subunits and a "master" regulatory protein termed NEMO (NF-κB essential modulator) or IKKγ. In presence of an immunostimulation, the IκB kinase phosphorylates two serine residues located in the SRD regulatory domain of IκB. When phosphorylated on these serines (e.g., serines 32 and 36 in human IκBα), the IκB proteins are modified by a process called ubiquitination, which then leads the degradation of IkB by a cell structure called the proteasome. With the degradation of IκB, the NF-κB complex is then free to enter the nucleus where it is able to 'turn on' the expression of the antimicrobial peptides and the following activation of the immune system.

Since data from scientific literature, (Falabella et al., 2018), seem to demonstrate that *Tn*BVANK1 is involved in the NF - kB signaling pathway, it was decided to express *Tn*BVANK1 in a model organism such as *Drosophila melanogaster*, in order to perform a functional analysis of this gene.

1.9 *Drosophila melanogaster* as a model system

The fact that 6 Nobel prizes were won by scientists who have used *Drosophila* as a model is only a confirmation of the importance that this organism has in the scientific field.
Since its first application for studying genetics performed by Thomas Hunt Morgan, *Drosophila* has been used to answer many questions about normal and pathological conditions. The major advantage of using the fruit fly is given by the multiple powerful genetic technique developed over the years, from gene knock-out to gene over-expression, that permit different analyses approaches within a tissue, or a cluster of cells, in the whole organism, in the larval development or in a defined time window. The high amenability of this organism is accompanied by an easier way to do it and to analyze the results, lower genetic redundancy and *low-cost* trials, that have no comparison with the other model systems used. This explains also why *Drosophila* is often exploited for genetic screens that have the purpose to find out the function of countless genes and to uncover the mechanism that lay behind a certain developmental process (St. Johnston 2002).

The fruit fly has a life cycle characterised by four phases: the embryo, the larva, the pupa, and the adult. Different tissues and different organs, as well as the entire organism, can be manipulated and used to study genes and biological processes. In this project I used this model organism and the TARGET system to express ubiquitously *Tn*BVANK1 in adult flies, both males and females, to evaluate the phenotypic effect of this viral protein and to understand its molecular mechanism of action.

1.10 *Tn*BVANK1 expression in *Drosophila melanogaster*

The expression of *Tn*BVANK1 in *Drosophila* started with the *Tn*BVANK1 expression in the *Drosophila* germline, which affects the localization of *gurken bicoid*, and *oskar*, mRNAs whose gene products are relevant for proper embryonic patterning (Duchi et al., 2010). The altered localization of these mRNAs is suggestive of general defects in the intracellular, microtubule-based, trafficking routes. Analysis of microtubule motor proteins components such as the dynein heavy chain and the kinesin heavy chain revealed defects in the polarized microtubule network. Interestingly, the *Tn*BVANK1 viral protein is uniformly distributed over the entire oocyte cortex and appears to be anchored to the microtubule ends. These data open up a very interesting issue on novel function(s) played by the *ank* gene family by interfering with cytoskeleton organization (Duchi et al., 2010).

Another research found that *Tn*BVANK1 expression during the development of *Drosophila,* more specifically in the larva prothoracic gland, blocks the larval-pupal molt. This phenotype can be rescued by feeding the larvae with 20-hydroxyecdysone. The localization of the *Tn*BVANK1 is restricted to the cytoplasm where it interacts with Hrs and Alix marked endosomes (Valzania et al; 2014). Collectively, these data demonstrate that the *Tn*BVANK1 protein acts as a virulence factor that causes the disruption of ecdysone biosynthesis and developmental arrest by impairing the vesicular traffic of ecdysteroid precursors in the prothoracic gland steroidogenic cells.

Based on these two last studies, my PhD project aims to ubiquitously express this protein throughout the entire *Drosophila* organism in order to evaluate its phenotypic effect also in the adult organism, to analyze both its molecular mechanism of action, and if *Tn*BVANK1 can be used as a bioinsecticide environmentally eco – sustainable.

2. MATERIALS AND METHODS

2.1 Fly strains

The following stocks were raised on standard cornmeal/yeast/agar medium at 21°C and crosses were made at 25°C unless otherwise stated.

1. The UASp-*TnBVank1* strain (genotype: *UASp-TnBVank1/UASp - TnBVank1; UASp-TnBVank1/UASp - TnBVank1; +/+*) was generated in our laboratory and it has 2 copies of *TnBVAnk1* gene in the X – chromosome and in the second chromosome.
2. The driver line used for the TARGET system (genotype: *w-/Y; TubGal80(Ts)/TubGal80(Ts); TubGal4/TM3, Ser*) was generated in our laboratory.
3. This last stock has been used as a control in this project #35787 of Bloomington Stock Center (genotype: *y,sc,v,sev/Y; +/+, UAS – mCherry/UAS – mCherry).*

2.2 TARGET system

The TARGET system, (temporal and regional gene expression targeting), (Mcguire et al., 2003), allows tissue-specific expression of a gene in a specific time interval of interest. By exploiting this strategy, it is

possible to avoid developmental arrest due to the *Tn*BVANK1 expression in the larval prothoracic gland (Valzania et al., 2014), allowing larval development to continue until reaching the adult stage. This makes it possible to induce the expression of *Tn*BVANK1 only in the adult organism.

A summary of the system used is illustrated in the following image (Fig 9.).

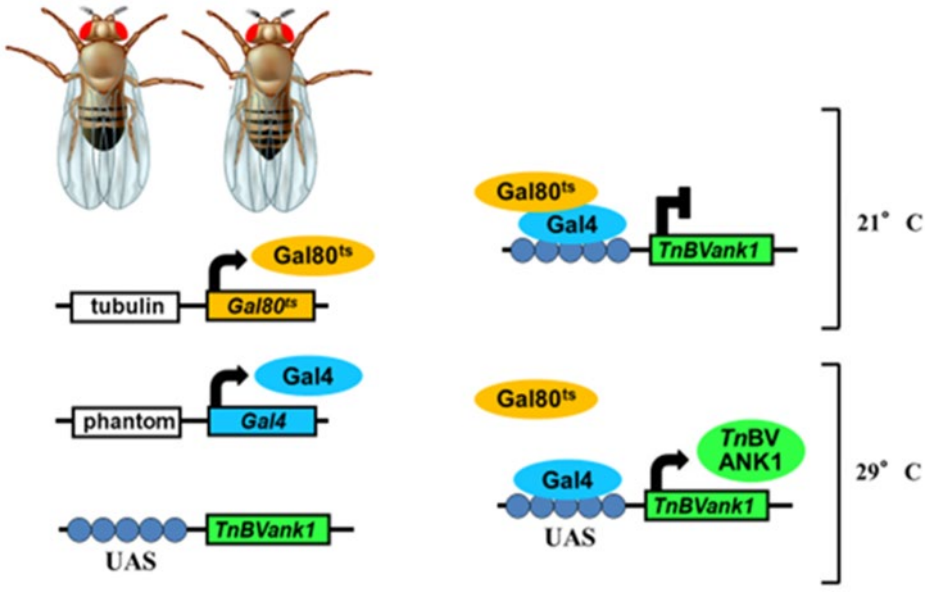

Figure 9: Overview of the TARGET system.

More specifically, this system involves using two *Drosophila* lines. A first line, called the 'responder line', contains the *TnBVAnk1* sequence within its genome, downstream of a UAS sequence.

A second line, called the 'driver line', contains the yeast transcriptional activator Gal4 sequence within its genome, downstream of a tissue-specific promoter. In this case, since the goal was to ubiquitously express *Tn*BVANK1, I used the promoter of *tubulin*, an ubiquitous housekeeping gene. Furthermore, downstream of the driver line's tissue-specific promoter is the sequence of a temperature-sensitive Gal4 inhibitor, known as Gal80(Ts). This system allows the control of Gal4 protein activity using a temperature-sensitive form of the Gal80(Ts) inhibitor. In this way, it is possible to control the expression of *TnBVAnk1* transgene by varying the growth temperature of *Drosophila melanogaster.*

By crossing the driver and the responder line, an F1 progeny is obtained, and that was allowed to develop at 21°C and moved to 29°C upon reaching the adult stage.

Thus, at 21°C, Gal80(Ts) binds Gal4, preventing it from inducing the transcription of *Tn*BVANK1. However, at the permissive temperature of 29°C, Gal80(Ts) is degraded, leaving Gal4 free to induce transcription and the resulting expression of *Tn*BVANK1.

2.3 Survival test

In order to perform the survival test, the following two strains (UASp-*TnBVank1* strain and the strain #35787 from Bloomington Stock Center) were left to grow at 25°C until a number of 150 male and female individuals for each strain were obtained. Furthermore, the following crosses show respectively the experimental and the controls individuals that were made using the TARGET system:

Parental males		**Parental virgin females**
w-/Y *TubGal80(Ts)/TubGal80(Ts)* *TM3Ser/TubGal4*	X	*yw, UAS-TnBVAnk1/yw, UAS-TnBVAnk1* *UAS-TnBVAnk1/UAS-TnBVAnk1* *+/+*

From the F1 progeny, 100 flies both males and females with the following genotype were selected:

- Females: *yw, UAS-TnBVAnk1/w⁻; TubGal80(Ts)/UAS-TnBVAnk1; TubGal4/+*
- Males: *yw, UAS-TnBVAnk1/Y; TubGal80(Ts)/UAS-TnBVAnk1; TubGal4/+*

This cross was established at the permissive temperature of 21°C, and the progeny were allowed to develop into adults. At this temperature, the inhibitor Gal80(Ts) inhibits Gal4 activity, blocking *TnBVAnk1* transcription.

By subsequently transferring the F1 progeny to a temperature of 29°C, Gal80(Ts) function is inhibited, allowing Gal4 to promote *TnBVAnk1* transcription.

The survival test was performed by monitoring the number of dead individuals at three- to four-day intervals.

As negative control for this survival test, individuals from the following two crosses were analyzed:

1)

Parental males		**Parental virgin females**
y,sc,v,sev/Y *+/+* *UAS-mCherry/UAS-mCherry*	X	*yw, UAS-TnBVAnk1/yw, UAS-TnBVAnk1* *UAS-TnBVAnk1/UAS-TnBVAnk1* *+/+*

From the F1 progeny, females and males with the following genotype were selected:

→ Females: *yw, UAS-TnBVAnk1/y,sc,v,sev; UAS-TnBVAnk1/+; UAS-mCherry/+*

→ Males: *yw, UAS-TnBVAnk1/Y ; UAS-TnBVAnk1/+; UAS-mCherry/+*

To strengthen the experimental data, I also used a second control cross in which male and female individuals were selected.

These adults carried the constructs for the expression of the transcriptional activator *Gal4*, its inhibitor *Gal80(Ts)*, and for *UAS-mCherry* .

2) **Parental males** X **Parental virgin females**

w^-/Y ;
TubGal80(Ts)/TubGal80(Ts) ;
TM3Ser/TubGal4

y, sc, v, sev / y, sc, v, sev ; /+/+ ;
UAS-mCherry/UAS-mCherry

From the F1 progeny, females and males with the following genotype were selected:

→ Females: *y, sc, v, sev / w^- ; TubGal80(Ts)/+; UAS-mCherry/TubGal4*

→ Males: *y, sc, v, sev / Y ; TubGal80(Ts)/+; UAS-mCherry/TubGal4*

The F1 progeny obtained from these control crosses, were born at 21°C, and then transferred to 29°C for the previously described survival test. The survival curves of the controls and of the experimental individuals were statistically analyzed using Kaplan-Maier curves and and log-rank analysis. All data were analyzed using Prism software, v. 5.0c (GraphPad software; San Diego, California, USA).

2.4 RNA extraction from adult flies

Working on Ice and using a potter, I homogenized the flies in 150µL of TRI Reagent® (Sigma-Aldrich) responsible to maintain the integrity of nucleic acids and at the same time lysate the cell membranes.

After adding 30 µl of chloroform (Sigma-Aldrich), I vortexed the samples for 15'' and stored for 13' at room temperature.

After that, I centrifuged the resulting mixture at 12000 x g for 15' at 4°C. Centrifugation separates the mixture into 3 phases: a red organic phase (containing proteins), an interphase (containing DNA), and a colorless upper aqueous phase (containing RNA).

Next, I transferred the upper aqueous phase in a fresh eppendor tube and I added 1 volume of phenol/chloroform/isoamyl alcohol, (Sigma-Aldrich), shaked (vortex) for 15'' and the resulting mixture was centrifuged at 12000 x g for 10' at 4°C.

Subsequently, I transferred the aqueous phase to a new fresh eppendorf tube. To precipitate RNA, I added 75µL of isopropanol (Sigma-Aldrich) and 2µL of glycogen (Sigma-Aldrich). The resulting mixture briefly shaken and stored for 8' at room temperature.

Right away, the mixture was centrifuged at 12000 x g for 10' at 4°C degrees.

After the centrifuge, I removed the supernatant and washed the RNA pellet by adding slowly 150 µL of ethanol at 75%. Then, the samples was centrifuged at 12500 x g for 5' at 4°C degrees.

Next, I took out the supernatant and I dried the RNA pellet with a Speed – Vac for 1' – 3', without letting that the RNA pellet dry completely as this will greatly decrease its solubility.

Subsequently, the RNA pellet was resuspend in 180µL of H_2O milliQ, and the samples was shaked, spined and leaved them on ice for 8'.

The second step started with the addition of 180µL (1 volume) of phenol/chloroform/isoamyl alcohol (Sigma-Aldrich). Shake vigorously and the centrifuge the resulting mixture at 12000 x g for 10' at 4°C.

After the centrifugation, transfer the aqueous phase in to a fresh eppendorf tube to precipitate the RNA with sodium acetate 3M (pH: 5,2). In the new eppendorf tube I added also 2µL of glycogen, (Sigma-Aldrich), 20µL of sodium acetate 3M (pH: 5,2) and 600µL (3 volumes) of EtOH 99%. Shake and store the samples at -20°C for 2 hours or at -80°C for 1,5 hour.

After the incubation, I centrifuged the samples at 14000 x g for 12' at 4°C. Then, I removed the supernatant and I washed the RNA pellet by adding slowly 200µL of EtOH 70%. Then, I centrifuged at 14000 x g for 5' at 4°C.

Subsequently, I removed the supernatant and I dried again the RNA pellet with Speed – Vac for 1' – 3'.

Right away, I resuspended the RNA pellet with 50µL of H_2O milliQ, shaking slowly and letting the samples in ice for 13'.

Finally, I read the RNA concentration and the values of 260 nm and the ratios 260 nm/280 nm and 260 nm /230 nm using the Nanodrope instrument (Eppendorf D30 BioPhotometer Photometer Xenon Micro

Nanodrop Measuring). I used 2µL of H_2O milliQ for the blank and 2µL of each samples -for the reading at the spectrophotometer.

2.5 Reverse transcription

Thaw on ice 4x (gDNA wiper HiScript III RT SuperMix for qPCR +gDNA wiper, Vazyme), 5x HiScript III, (gDNA wiper HiScript III RT SuperMix for qPCR +gDNA wiper, Vazyme), and the RNA samples. Spin down, mix by flicking the tube, and then spin down again. In 0.2mL eppendorf tubes I added the following reagents:

1. 350 ng of RNA
2. 4µL of gDNA wiper mix (gDNA wiper HiScript III RT SuperMix for qPCR +gDNA wiper, Vazyme)
3. ddH_2O to bring up to a volume of 12µL

Next, I incubated the samples in the thermocycler at 42°C for 2 minutes.

During this incubation, the DNase contained in the gDNA wiper mix degrades any impure DNA molecules present in the extracted RNA sample.

Subsequently, I added in each 0,2mL eppendorf tube 4µL of HiScript III and incubate in the thermocycler with the following program:

1. 37°C for 15'
2. 85°C for 5"

Finally, in order to perform a qPCR – assays, it is necessary to store the cDNA obtained at least 1 night at –20°C.

2.6 qPCR analysis

The Real-Time PCR allows the quantification of a transcribed gene by evaluating the amplification of its cDNA. To evaluate the differences existing between the samples, a housekeeping gene is used as a "reference", because its level of expression is presumably constant even between different samples.

The reference gene is the housekeeping *RPL32*, which encodes for ribosomal protein that is a component of the 60S subunit. The protein belongs to the L32E family of ribosomal proteins.

To perform this qPCR assay, first of all, it is necessary to thaw in ice the reaction mix (SsoAdvanced Universal SYBR Green Supermix, BioRad) and the necessary primers.

The SYBR green is an intercalant that emits fluorescence while binding the minor DNA groove; the number of fluorescent molecules increase along with the number of copies of the amplicon, showing the quantity of each target.

The qRT-PCR analysis was performed considering 3 technical replicate, for each biological replicate. The volume inside of each well of the plate was 10µL:

1. 5µL of SsoAdvanced Universal SYBR Green Supermix
2. 3µL of ddH_2O
3. 1µL of the cDNA

For each biological replicate, it is necessary to calculate the number of wells of the plate that are necessary for the analysis, and then, prepare the sample to subsequently aliquoted into each well of the plate. Once all samples have been prepared, aliquot 9µL of each sample into each corresponding well. Finally, complete the qRT - PCR plate setup by adding 1µL of the Fw and Rv primer mix to each well corresponding to the gene that you want to analyze. The two primers inside the mix have a concentration of 2.5µM each.

Finally, spin the plate and put it into a thermocycler (CFX96™ System, Biorad) with the following cycle:

1. 95°C x 30”
2. 95°C x 10” ┐
3. 60°C x 20” ├ 40X
4. 65°C x 5” ┘
5. 95°C x 5”

Once qPCR is complete, the output is C_q values. C_q (quantification cycle) values in real-time PCR indicate the number of amplification cycles required for the fluorescence produced by the reaction to exceed a predetermined threshold. A low C_q value means that the fluorescence threshold was reached in just a few cycles, suggesting a higher initial concentration of the target genetic material in the sample. Conversely, a high C_q value indicates that it took many cycles to reach the threshold, indicating a lower initial amount of genetic material. First of all, the C_q values obtained from the three technical replicates for each biological replicate of the gene to be quantified and of the reference gene (*RPL32*) are averaged. Subsequently, for each biological replicate, the average of the C_q of the three technical replicates of the reference gene is subtracted

from the average of the C_q of the three technical replicates of the gene to be quantified, thus obtaining the ΔC_q value.

$\Delta C_q = C_{q(target)} - C_{q(\textit{RPL32})}$

This normalizes the expression of the target gene relative to that of the reference gene.

Next, the 'calibrator' is calculated, which means the arithmetic mean of the ΔCq values of the control samples of the gene being quantified, and then the calibrator is subtracted from all previously obtained ΔCq values. This step allows us to compare gene expression between the two groups, the experiment and the control for the gene being quantified.

$\Delta\Delta C_q = \Delta C_q$ – 'calibrator value'

Finally, the $2^{\wedge -\Delta\Delta Cq}$ is calculated for each biological replicate. This will give you the fold change in relative gene expression for that sample compared to the control. The fold change data therefore represents, for each biological replicate, the quantification of the transcript of the gene being quantified.

The following table (Table 1) shows the primer sequences used for each gene analyzed.

Target genes	**Primer sequence 5' → 3'**
Rpl32_FW	ATCGGTTACGGATCGAACAA
Rpl32_RV	GACAATCTCCTTGCGCTTCT
Drs_FW	CGTGAGAACCTTTTCCAATATGATG
Drs_RV	TCCCAGGACCACCAGCAT
Dpt_FW	GCTGCGCAATCGCTTCTACT
Dpt_RV	TGGTGGAGTGGGCTTCATG
Att.2_FW	CCTTGACGCACAGCAACTTC
Att.2_RV	CCGATCCCGTGAGATCCAAG
Def_FW	GTTCTTCGTTCTCGTGG
Def_RV	CTTTGAACCCCTTGGC
Dro_FW	TTTTCCTGCTGCTTGCTTGC
Dro_RV	GGCAGCTTGAGTCAGGTGAT
Gal4_FW	GTGTGCAATCCCATTACCGC
Gal4_RV	CACAGGTGAAGGCCCTACTG
*TnBVAnk1*_FW	AATGCACCCAACCAAACT
*TnBVAnk1*_RV	AGCACAGCCATTTCGCCA

Table 1: Target genes and primer sequences used for qRT-PCR analysis. RPL32 is used as housekeeping gene.

2.7 Expression of *Tn*BVANK1

In order to express *Tn*BVANK1 in *E.coli* BL21, I first designed and purchased from the GeneScript company, the following plasmid, pET - 11a (Fig. 10) in which the *Tn*BVANK1 sequence was inserted in frame at the 3' end into the MCS with a His - Tag composed of 6 Histidines.

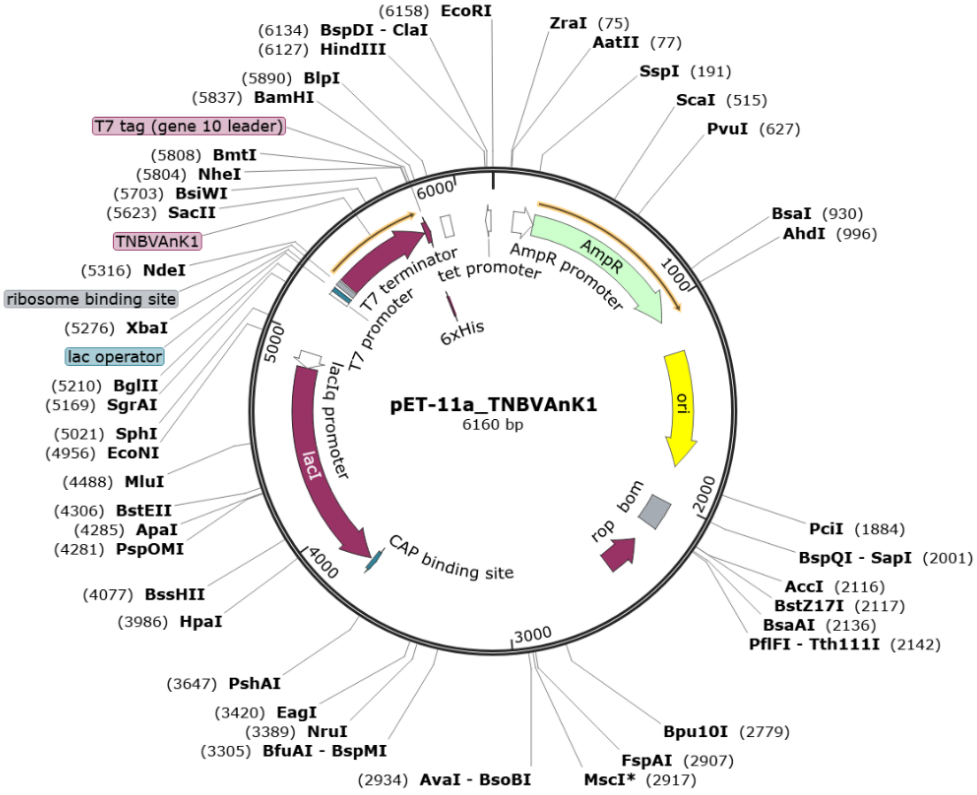

Figure 10: E.coli expression vector used to express TnBVANK1 using the inducible lac promoter.

To amplify the resulting plasmid, the *E. coli* DH5α strain was transformed and then extracted using mini-prep. At this point, the plasmid was double-digested using NdeI and BamHI, checked for the presence of the *Tn*BVANK1 sequence using an agarose gel, and finally the extracted plasmid was sent for sequencing. Sanger sequencing provided final confirmation that the *Tn*BVANK1 gene sequence was correctly included in the plasmid extracted.

At this point, I prepared 4 liters of sterile minimal media, (1xMSM, Sigma Aldrich), supplemented with 1X ampicillin (Sigma Aldrich) to grow 4 liters of *E.coli* BL21 at 22.5°C with 250 rpm until OD 0.8-0.9.

After that, I induced the expression with 1mM of IPTG at 22°C for 21 hours at 250 rpm. Once the 21 hours were concluded, the 4 liters of culture were centrifuged using the Himac centrifuge, High-speed Refrigerated

Centrifuge model CR22N, at 7000 rpm 4°C, 15'. Subsequently, the pellet was separated from the supernatant, weighed, and resuspended in the same equilibration buffer as the chromatography column (HisPrep FF 16/10) for affinity chromatography. It was observed that the optimal ratio between buffer volume per gram of pellet was 50 mL per 5 g of pellet. The buffer composition is as follows: Buffer A equilibration: 50 mM phosphate buffer, 500 mM NaCl, 5 mM imidazole pH 7.4.

Once resuspended, a cocktail of protease inhibitors (Sigma Aldrich) and lysozyme (Sigma Aldrich), are added to the pellet, and the the pellet was lysed by sonication with the SONOPULS HD 5020 sonicator. The sonication consisted of four 30-second cycles with a 1-second pulse on/1-second pulse off rhythm and an AM of 60%.

At this point, the resulting mixture is centrifuged at 20,000 rpm at 4°C for 45 minutes. This step separates the cellular debris from the protein lysate present in the supernatant. Once this step is complete, the pellet is discarded and the protein lysate is added to the chromatographic column for affinity chromatography.

At this point, the resulting mixture is centrifuged at 20,000 rpm at 4°C for 45 minutes. This step separates the cellular debris from the protein lysate present in the supernatant. Once this step is complete, the pellet is discarded and the protein lysate is added to the chromatographic column for affinity chromatography to purify my protein of interest, *Tn*BVANK1.

2.8 Purification of *Tn*BVANK1

The chromatographic column used for affinity chromatography and the instrument used are: HisPrep FF 16/10 with a volume capacity of 10 mL (Cytiva) and ÄKTA pure (Cytiva). Before adding the protein lysate, the column is equilibrated at a flow rate of 2.5 mL/min with 50 mL of equilibration buffer A with the following composition: 50 mM phosphate buffer, 500 mM NaCl, 5 mM imidazole.

A small concentration of imidazole is necessary to avoid nonspecific interactions with the chromatographic column. The protein lysate is then loaded at a flow rate of 2.5 mL/min and finally eluted with a 0-100% gradient in 2 volumes of buffer B (50 mM sodium phosphate, 500 mM NaCl, 500 mM imidazole) at a flow rate of 1 mL/min. Fractions were collected manually following the UV absorbance peak. The entire process is performed at 4°C in a cold room.

At this point, the fractions corresponding to the affinity chromatography elution peak are loaded into the chromatographic column for size exclusion chromatography (SEC) the same day. Size exclusion chromatography allows for better purification of *Tn*BVANK1 by eluting away high-molecular-weight proteins. The instrument used for this analysis is again ÄKTA pure, while the column used are the following: HiLoad Superdex 16/600 75 pg or HiLoad Superdex 16/600 200 pg. Both columns are optimal for the

purification of *Tn*BVANK1 (18kDa) since they have a pore size that allows the separation of proteins with a molecular weight between 3kDa-70kDa and between 10kDa-600kDa, respectively.

In the SEC, the sample is placed in a 50 mL "Superloop" previously rinsed with water, and then equilibrated with Buffer B. In this case, equilibration is performed with Buffer B because the fractions eluted from the affinity chromatography were eluted with Buffer B. After equilibration, the corresponding fractions are placed in the "Superloop," and finally the "Superloop" is connected to the ÄKTA pure instrument. The column is then connected and equilibrated with 150 mL of Buffer B with a flow rate of 1.0 mL/min. The sample application and the elution were performed at a flow rate of 1 mL/min. The entire process also for the size exclusion chromatography is performed again at 4°C in cold room.

2.9 SDS-PAGE assay

Once the SEC is completed, the fractions corresponding to the elution peaks are run in an SDS-page assay to verify the presence of my protein of interest, *Tn*BVANK1. For this analysis, a 12.5% polyacrylamide gel was prepared with a 5% stacking gel. The mixture required for the 12.5% stacking gel is as follows:

1. 30% acrylamide solution: 3,3 ml
2. Separating gel buffer: 1,8 ml
3. H_2O: 2,6 ml
4. TEMED: 7,5 µl
5. Ammonium persulfate (APS) (10%): 45 µl

The separating gel buffer has the following composition: 1.5M Tris pH 8.8, SDS 0.4%

The mixture required for the 5% stacking gel is as follows:

1. 30% acrylamide solution: 0,7 ml
2. 5% stacking gel buffer: 1,2 ml
3. H_2O: 2,7 ml
4. TEMED: 4,5 µL
5. APS (10%) 45 µL

The gel stacking buffer has the following composition: 0.5 M Tris pH 6.8 SDS 0.4%

Once the gels are prepared, the samples to be loaded are prepared by also adding the running buffer (Laemmli SDS sample buffer, reducer (4X) 25 mL Thermofisher Scientific), then incubating the samples at 70°C for 5 minutes.

Once the polyacrylamide gel has solidified, you can load 4µL of the marker (Unstained Protein Molecular Weight Marker, Thermofisher Scientific) and 10µL of the pre-prepared samples into the wells. Finally, add the running buffer (25 mM Tris, 250 mM glycine, 0.1% SDS) and set the power supply to 90V until the samples reach the separating gel. At this point, increase the voltage to 130V for one hour.

Subsequently, at the end of the electrophoretic run, the gel is incubated inside a tray with the following solution (Thermo Scientific™ GelCode™ Blue Stain Reagent) with the aim of carrying out a quick staining and verifying in which fractions my protein of interest *Tn*BVANK1 was present. At this point, the fractions containing *Tn*BVANK1 are combined and the sample concentration is calculated using UV-visible spectrometry. The combined fractions are then dialyzed in two steps.

1) A dialysis solution containing 50 mM phosphate buffer, 300 mM NaCl, and 300 mM imidazole is prepared.

2) The second dialysis step involves the use of a solution composed of 50 mM phosphate buffer, 150 mM NaCl, and 150 mM imidazole.

Finally, the sample concentration is increased to 150 mM using a centricon tube (Sigma-Aldrich) and the sample was stored at 4°C in a cold room waiting for the analysis.

2.10 CD – spectroscopy analysis

The CD spectrum was obtained using a J-1500 CD Spectrometer (Jasco), a circular dichroism spectrometer. The quartz cuvette used was 12.5 x 12.5 x 45 mm, with a 10 mm optical path and a sample volume capacity of 3.5 mL.

Our goal with the CD spectrum was to show the percentage of different secondary structures (alpha helices and beta sheets) present in the *Tn*BVANK1 protein. Therefore, a wavelength range of 195 nm to 250 nm was chosen (both for the sample and the control spectra), which includes the absorption of the peptide bond at 220 nm.

Furthermore, since the protein of interest *Tn*BVANK1 is concentrated 150mM and it is dissolved in a buffer composed of 50mM sodium phosphate, 150mM imidazole and 150mM NaCl at pH 7.2, it is necessary to perform a control spectra with only the buffer in order to verify that at that concentration of imidazole, its adsorption, (at 210-215 nm and with a low intensity peak at 280 nm), was so negligible as not to alter the absorption of the protein of interest *Tn*BVANK1. The protocol begins by cleaning the quartz cuvette with 70% Et-OH and distilled water, then adding 3.5 mL of buffer only and observing the average spectrum after 185 acquisitions at a scanning speed of 100 nm/min. Once it is confirmed that the imidazole absorption is

negligible at that concentration, the same protocol is repeated to obtain the spectrum of the *Tn*BVANK1 sample.

2.11 NMR analysis

The NMR analysis of an ^{15}N-^{1}H HSQC spectrum was performed on a BRUKER 800 MHz proton frequency instrument equipped with a cryoprobe. The measurement times were set up to 6 hours.

The NMR tubes used are 5 mm diameter borosilicate glass tubes containing a sample volume of 550 µL plus 50 µL of deuterated water (D_2O). The deuterated water is necessary to prevent the water's protons from interfering with the signal of the proton ^{1}H in the protein of interest *Tn*BVANK1. Deuterium (D), an isotope of hydrogen, is not detected in proton NMR spectroscopy, thus allowing for sharper and cleaner spectra. Once the NMR sample has been prepared, it is stored in a cold room at 4°C while waiting for the measurement.

3. RESULTS

3.1 The effect of *Tn*BVANK1 ubiquitous expression on the survival of adult flies

With the aim to analyze the effect of the *Tn*BVANK1 ubiquitous expression on the survival of adult male and female individuals expressing the transgene, I performed a survival test (Fig.11A, 11B).

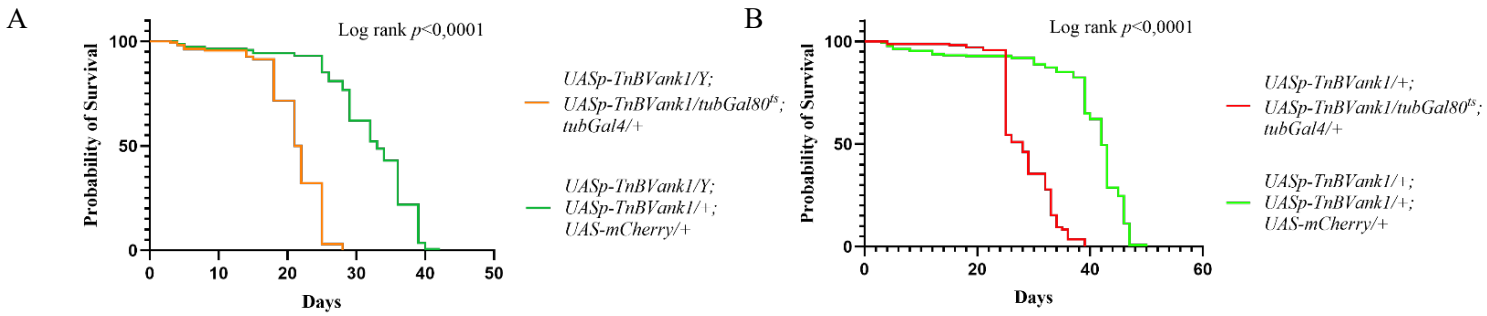

Figure11: These two graphs respectively represent the results of the survival test in females (A) and males (B) that were selected from F1 after crossing using the TARGET system. Both Kaplan-Maier plots show a statistical analysis of Log rank < 0.0001. This means that the control and experimental survival curves are statistically different.

ubiquitously expressing *Tn*BVANK1 show a peak in mortality 24 days after the transfer to 29°C and the beginning of the transgene expression compared to their control counterparts, which show a peak in mortality only after 42 days (Fig. 11A). Female transgenic individuals exhibit complete mortality after 39 days, compared to control individuals who exhibit mortality after 50 days. Male individuals, show a first peak in mortality 14 days after *Tn*BVANK1 expression compared to control individuals, which showed the

peak of mortality after 27 days (Fig. 11B). Transgenic males exhibit complete mortality after 28 days, compared to control individuals, which exhibit mortality after 42 days.

Taken together, these data shown that ubiquitous expression of *Tn*BVANK1 in *Drosophila* adult male and female causes a phenotype of reduced viability and consequent premature death compared to control individuals.

3.2 Ubiquitous expression of *Tn*BVANK1 in adult *Drosophila melanogaster* induces an increased of transcript levels of five antimicrobial peptides genes.

The following papers Badinloo et al., 2018, Kounatidis et al., 2017 and Cao et al., 2013, show that the overexpression of antimicrobial peptides (AMPs) in *Drosophila melanogaster* leads to marked death/shortening of adult life. In particular, the paper published by (Badinloo et al., 2018), highlights that ubiquitous overexpression of several antimicrobial peptides induces tissue apoptosis, mitochondrial depolarization, and a significant reduction in adult lifespan. This is the most explicit demonstration of AMP toxicity in vivo.

Based on this literature, our hypothesis was that the premature death of individuals ubiquitously expressing *Tn*BVANK1 was the hyperactivation of the innate immune system and a consequent overexpression of AMPs genes. In order to support this thesis, I analyzed the expression of five AMPs in *Tn*BVANK1 transgenic flies and in the control. First of all, two time points were chosen: the first, after 4 days of expression, when the individuals were mostly still alive, and the second after 14 days of expression, when the individuals began to die (Fig. 12).

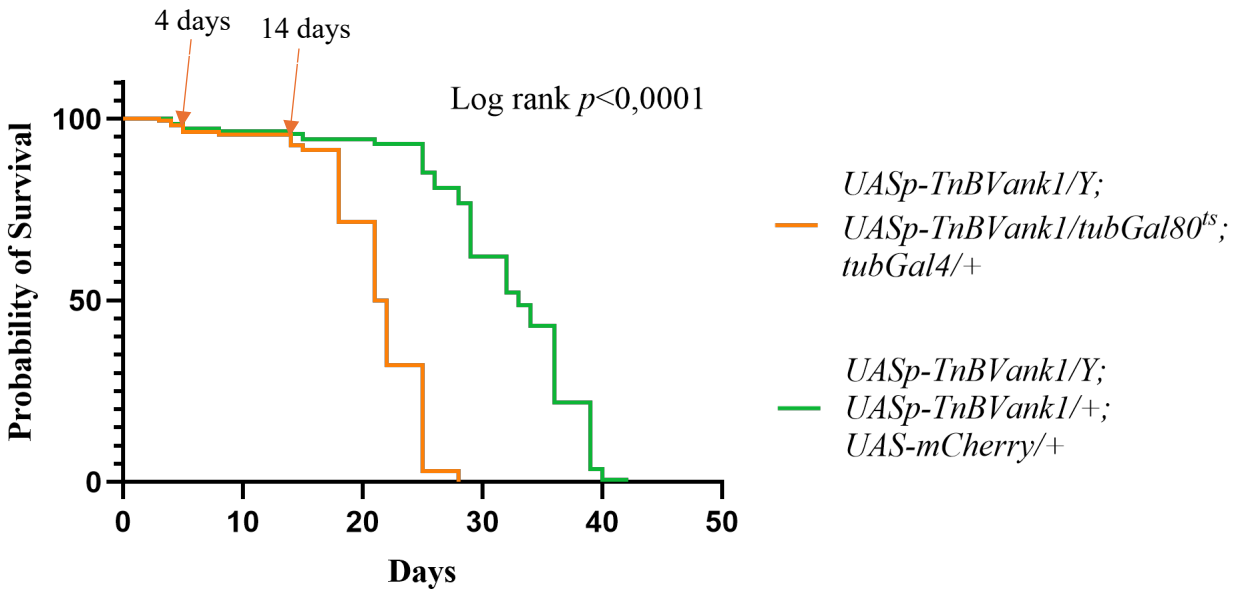

Figure12: Overview of the survival test in adult male flies following ubiquitous expression of TnBVANK1.

Before to start the expression analysis of the AMPs, I tested whether the expression level of the transgene was subjected to changes over time. I found that the *Tn*BVANK1 transcript level was significantly increased in 14 days adult transgenic individuals in comparison to 4 day old individuals (Fig. 13A). However, we decided to perform a second qRT-PCR to evaluate the Gal4 transcript level after 14 days of expression (Fig. 13B). This last control analysis was essential to verify that the increase in *Tn*BVANK1 transcript was not due to the TARGET expression system used.

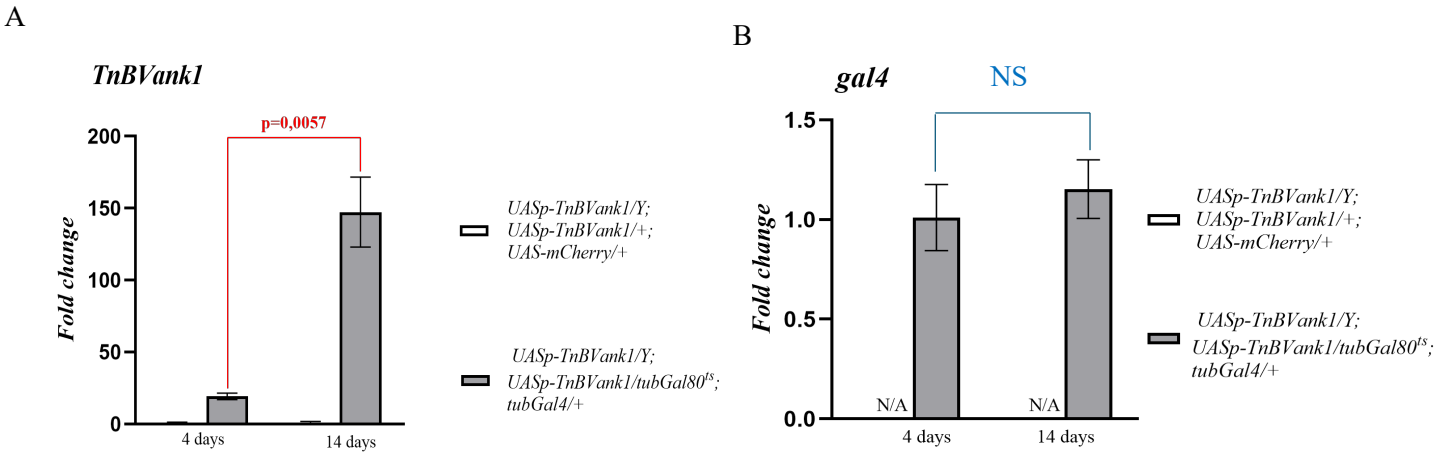

Figure13: These PCR plots show respectively the statistically significant increase in TnBVANK1 transcript after 14 days of expression (Fig.A), and the trend of Gal 4 transcript after 14 days of expression. (Fig.B). Statistical analysis was performed with the Student's T-test.

As can be seen in (Fig.13B), the Gal4 transcript level does not vary between 4 and 14 days of expression. This latter data confirms that the significant increase in *Tn*BVANK1 transcripts after 14 days of expression compared to 4 days of expression is not caused by the TARGET expression system used.

Right away, aiming to follow up on the hypothesis of premature death induced by an hyperactivation of the immune system with consequent overexpression of antimicrobial peptides, we decided to evaluate the expression levels of several antimicrobial peptides already studied in the article published by (Badinloo et al., 2018) comparing 4 and 14 days old individuals. The results of this qRT-PCR analysis are listed in (Fig. 14).

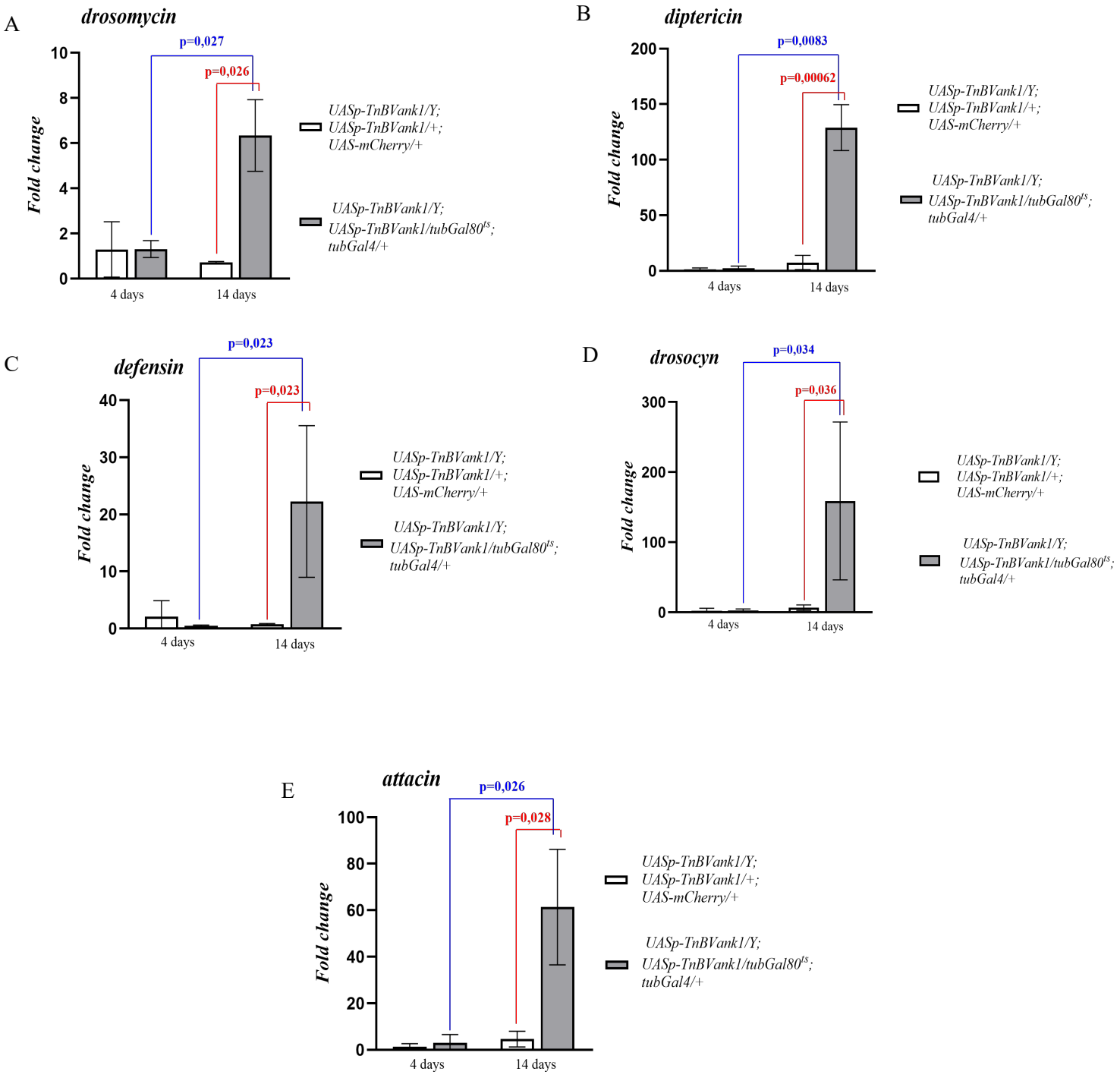

Figure14: These qPCR data show the increases after 14 days of ubiquitous expression of the transcript levels of five antimicrobial peptides: Drosomycin (A), Diptericin (B), Defensin (C), Drosocyn, (D) Attacin (E). The statistical analysis was performed with the Student's T-test.

These latest qRT-PCR data show how male flies ubiquitously expressing *Tn*BVANK1 exhibit a significant increase in the transcription level of different AMPs genes between 4 and 14 days, supporting the hypothesis of death due to an anomalous overexpression of several antimicrobial peptides.

3.3 *Tn*BVANK1 was correctly expressed in *E.coli* BL21

The presented data suggest that *Tn*BVANK1 has the potential to induce lethality in *Drosophila*. It is possible that this virulence factor could have similar effects also on other insects, opening the possibility to use it as an insecticide with low toxicity towards no target species. One feature of a candidate active molecule is the possibility of modifying its structure to improve its activity or modify its specificity. These manipulations require to know the protein tertiary structure in order to perform a single-aminoacids modification by rational protein engineering. To achieve this goal, this structural study aims to identify specific amino acid residues in the *Tn*BVANK1 which are responsible for its effect.

Therefore, since the tertiary structure of *Tn*BVANK1 has not yet been studied in scientific literature, it was decided to resolving its tertiary structure using circular dichroism spectroscopy (CD – spectroscopy) and nuclear magnetic resonance spectroscopy (NMR). To this aim, my work started by evaluating whether the expression of *Tn*BVANK1 in *E.coli* BL21 was successfully succeeded. (Fig. 15)

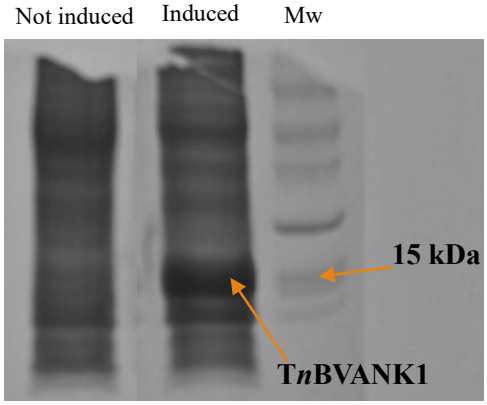

Figure 15: SDS-PAGE analysis showing the presence of the TnBVANK1 band starting from E.coli cells in which the expression was induced, compared to control E.coli cells in which the expression of TnBVANK1 was not induced.

The (Fig. 15) shows the SDS-PAGE analysis performed to demonstrate that *Tn*BVANK1 expression in *E.coli* BL21 was successful.

3.4 Purification of *Tn*BVANK1 by affinity chromatography (IMAC) and by size exclusion chromatography (SEC)

Once it was verified that the *Tn*BVANK1 expression protocol was successful, it is possible to proceed with the purification of *Tn*BVANK1 by the two chromatographies (Fig. 16).

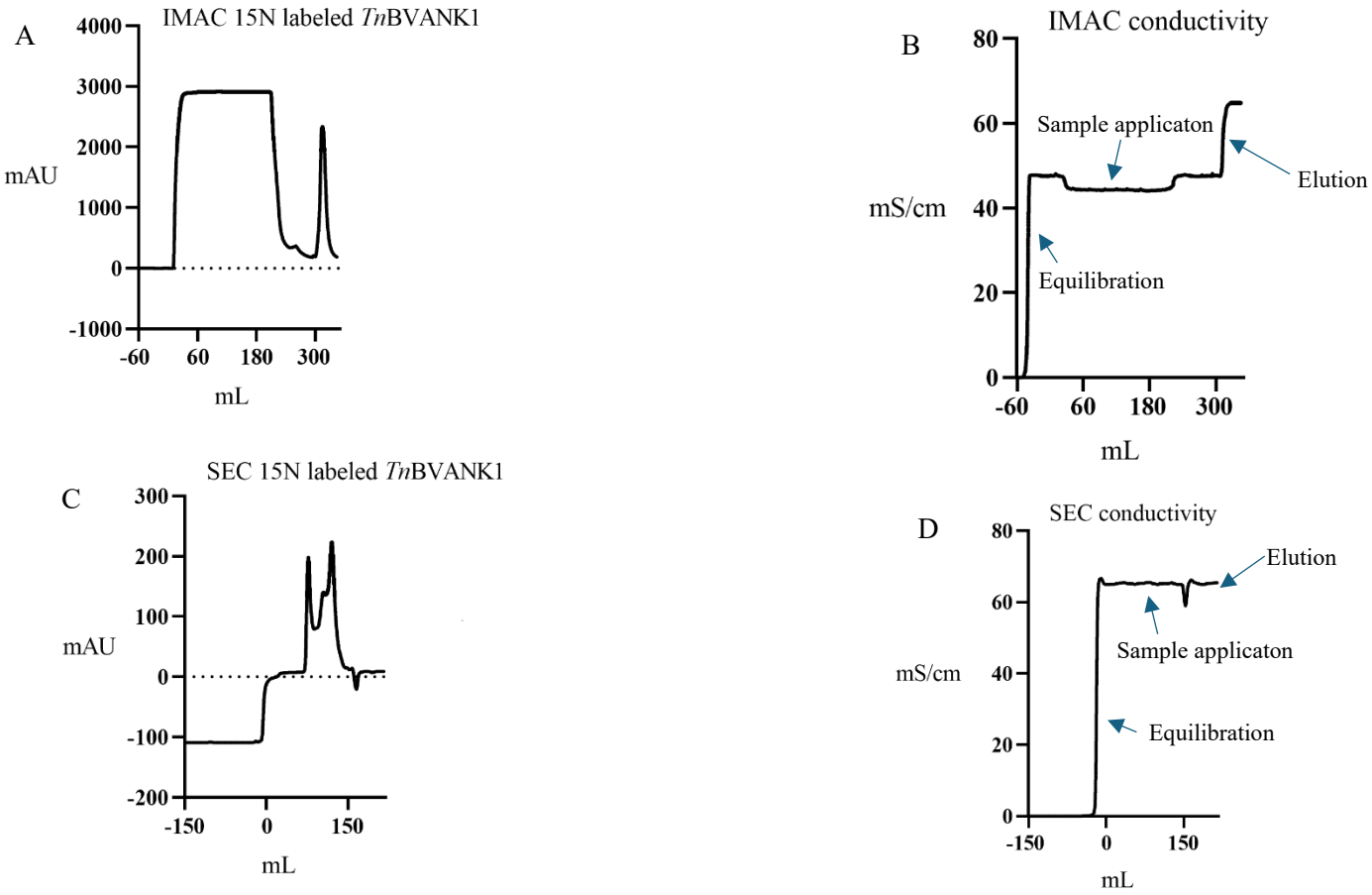

Figure16: The (Fig. 16A, 16C), show the two chromatograms of the IMAC and of the SEC, while (Fig. 16B, 16D) show the respective conductances of the two chromatographies. The chromatogram represents the UV absorption spectrum at 280 nm expressed in mAU, while the conductances is the ability to conduct electric current, which depends on the presence of free ions (electrolytes) in the buffer solution. The unit of measurement of conductance is expressed in mS/cm.

The IMAC (Fig. 16A) performed is a type of chromatography that exploits the specific binding between the Histidin tag fused to the C-terminus of *Tn*BVANK1 and the nickel present in the column used. Therefore, upon elution, an elution peak can be observed indicating that *Tn*BVANK1 has been eluted from the chromatographic column.

SEC (Fig. 16C), unlike IMAC, is a type of chromatography that separates proteins based on their size and molecular weight. In this case, three elution peaks can be observed, corresponding to elutions of proteins

with different sizes and molecular weights. In order to observe in which SEC peak *Tn*BVANK1 was eluted, the fractions corresponding to the three SEC elution peaks were analyzed by a second SDS-PAGE assay.

Finally, by observing the conductances of the two chromatographies (Fig. 16B, 16D), it is possible to understand that both chromatographic processes were carried out correctly. This is because in the IMAC, column equilibration and sample application are performed with the same buffer A, therefore, it is possible to observe how the conductance remains stable at the time of sample application. However, elution is performed with buffer B, which contains a higher concentration of imidazole than buffer A, therefore, in (Fig. 16B), an increase in conductance can be observed at the time of elution. Unlike the IMAC, in the SEC (Fig. 16D), column equilibration, sample application, and elution were performed with the same buffer B, therefore, it is possible to observe how after equilibration, the conductance remains stable.

3.5 Evaluation of *Tn*BVANK1 expression by SDS-PAGE and mass spectrometry analysis

The following image (Fig. 17) shows the fractions of the third peak of the size exclusion chromatography that were run in the SDS-PAGE assay.

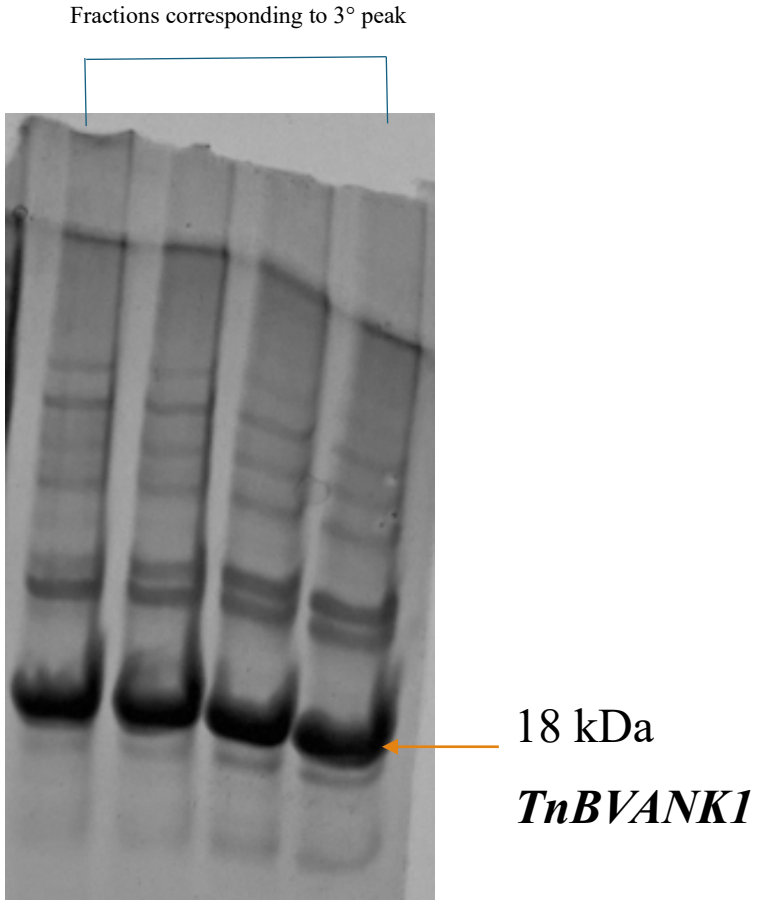

Figure17: SDS-PAGE analysis of the fractions concerning the third peak of the size exclusion chromatography. This assay highlight that TnBVANK1 is present with an high intensity in all the fractions of the third peak, but, 9 impure high-molecular-weight proteins with a low intensity are still present in the fractions.

The (Fig. 17) shows the SDS-PAGE assay performed to verify in which elution peak of the SEC *Tn*BVANK1 was eluted. The data showed that *Tn*BVANK1 was eluted in the third peak of the SEC.

At this point, we wanted to analyze the identity and type of these impure proteins, still present after size exclusion chromatography using mass spectrometry. This data will be listed in this table (Table 2).

Protein ID from Uniprot
Ank1 protein OS=Toxoneuron nigriceps polydnavirus
ECOLI 50S ribosomal protein L15 OS=Escherichia coli (strain BL21)
ECOLI 30S ribosomal protein S5 OS=Escherichia coli (strain BL21)
ECOLI Chaperone protein DnaK OS=Escherichia coli (strain BL21)
ECOLI 30S ribosomal protein S18 OS=Escherichia coli (strain BL21)
ECOLI 30S ribosomal protein S4 OS=Escherichia coli (strain BL21)
ECOLI 50S ribosomal protein L4 OS=Escherichia coli (strain BL21)
>ECOLI 30S ribosomal protein S13 OS=Escherichia coli (strain BL21)
ECOLI 50S ribosomal protein L17 OS=Escherichia coli (strain K12)

Table 2: Summary of the mass spectrometry results about all the protein that are present in the third fraction of the size exclusion chromatography.

As can be seen from the table, in the fractions relating to the third peak, the impure high molecular weight proteins correspond to ribosomal proteins and proteins of the chaperonin class.

The mass spectrometry data appears to tell us that *Tn*BVANK1 does not elute alone from the chromatographic column but elutes in interaction complexes with these impure proteins.

Despite mass spectrometry data demonstrating the presence of other impure proteins in the purified sample, the decision was anyway to perform a structural analysis of the sample by CD-spectroscopy and NMR.

3.6 Analysis of the structure of *Tn*BVANK1 by circular dichroism spectrometry (CD-spectroscopy) and nuclear magnetic resonance spectrometry (NMR)

Before the acquisition of the CD spectrum of the *Tn*BVANK1 sample, a control spectrum with buffer only was verified to ensure that the 150 mM imidazole concentration did not induce imidazole absorption in the CD spectrum (Data not shown).

Subsequently, once the CD spectrum of the sample was obtained, it was analyzed using Bestsel software, which provides the percentage of alpha helices, loops, beta sheets, and random coils present in a sample as output. The result is shown in (Fig. 18).

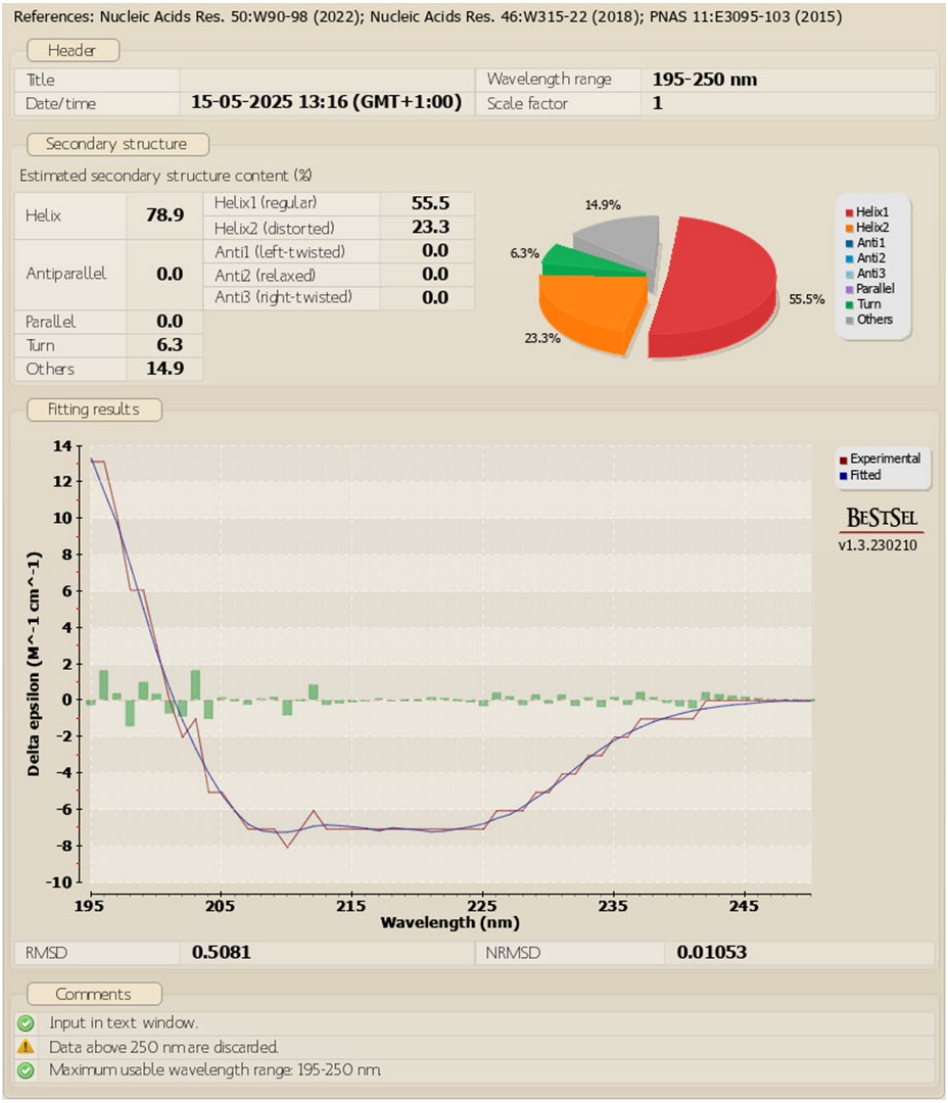

Figure 18: The figure shows the CD spectrum results analyzed with the Bestsel software.

The CD spectrum of *Tn*BVANK1 was calculated over a wavelength range from 195 to 250 nm. This range is necessary for studying the protein's secondary structure since peptide bonds absorb at shorter wavelengths, such as around 220 nm. As can be seen in Figure 15, the red curve is the CD spectrum of the sample of interest, *Tn*BVANK1, while the blue curve is the predicted and optimal curve for the obtained $\Delta\varepsilon$ values. $\Delta\varepsilon$ represents the differential molar absorbance between left (L) and right (R) circularly polarized light for a chiral molecule, calculated as $\Delta\varepsilon = \varepsilon_L - \varepsilon_R$. It is a measure of the difference in absorption of these two types of polarized light at a specific wavelength. Furthermore, the RMSD values are 0.50 and the NMRSD is 0.01. The RMSD measures the absolute difference between the measured and predicted spectra, respectively. In this case, a lower RMSD indicates a better agreement, but this is directly influenced by the intensity of the CD signals, so a higher intensity spectrum will naturally have a higher RMSD. Therefore, the NMRSD is needed which measure the normalized root mean square deviation, a statistical measure of the difference

between an experimental and a calculated circular dichroism (CD spectrum) of a protein, used to evaluate the accuracy of secondary structure prediction from CD data. A low NMRSD value indicates that the predicted and experimental spectra are similar, suggesting a good fit and accurate structural information, while a high NMRSD value suggests a poor fit and less reliable structural data. In NMRSD, unlike RMSD, the differences are "normalized" by the CD signal intensity to provide a relative measure of error, often used in bioinformatics and structural analysis. An NMRSD of a CD spectrum considered optimal for secondary structure analysis is <0.1, therefore, since the NMRSD obtained for the *Tn*BVANK1 sample is 0.01, this value can be considered optimal. The NMRSD value demonstrates that the circular dichroism spectrum was obtained and performed correctly. Finally, the results show that the *Tn*BVANK1 sample displays 78.9% alpha helices, 6.3% loops, and 14.9% random coils. Specifically, regarding the alpha helices, 55.5% are regular alpha helices and 23.3% are distorted alpha helices. These results seem to indicate that the *Tn*BVANK1 protein is not perfectly folded since it has a 14.9% of random coils. However, in the section discussing the results, we will try to provide an explanation for this finding.

Subsequently to the CD spectrum, with the aim of analyzing even better and more completely the tertiary structure of *Tn*BVANK1, it was decided to carry out a more complete structural analysis by nuclear magnetic resonance (NMR) spectroscopy with a ^{15}N-^{1}H HSQC spectra.

The heteronuclear quantum coherence or heteronuclear quantum correlation experiment, usually abbreviated as HSQC, is frequently used in NMR spectroscopy of organic molecules and is of particular importance in the field of protein NMR. The experiment was first described by Geoffrey Bodenhausen and D. J. Ruben in 1980. The resulting spectrum is two-dimensional (2D) with one axis for the proton (1H) and the other for a heteronucleus (an atomic nucleus other than a proton), which is 15N. The peculiarity of this spectrum is that it contains a peak for each unique 1H attached to 15N.

Every residue in the protein, except proline, has an amide proton bonded to a nitrogen atom in the peptide bond. HSQC provides the correlation between the nitrogen and the amide proton, and each atom produces a peak in the HSQC spectra. Each residue (except proline) can therefore produce an observable peak in the spectra,

The ^{15}N-^{1}H HSQC spectra is typically the first heteronuclear spectrum acquired for resonance assignment, in which each amide peak, and each side – chains peak is assigned to a specific residue in the protein. If the protein is folded, the peaks are usually well dispersed, and most individual peaks can be distinguished. However, the presence of a large cluster of strongly overlapping peaks around the central part of the spectrum indicates the presence of significant unstructured elements in the protein. In these cases, where strong resonance overlap occurs, assigning resonances in the spectra can be difficult.

As can be seen from, (Fig. 19), it was decided to perform two spectra: first, a spectrum without 5M UREA added to the sample, (Fig. 19A), and then a second spectrum with 5M UREA added (Fig. 19B).

A Without UREA 5M

B With UREA 5M

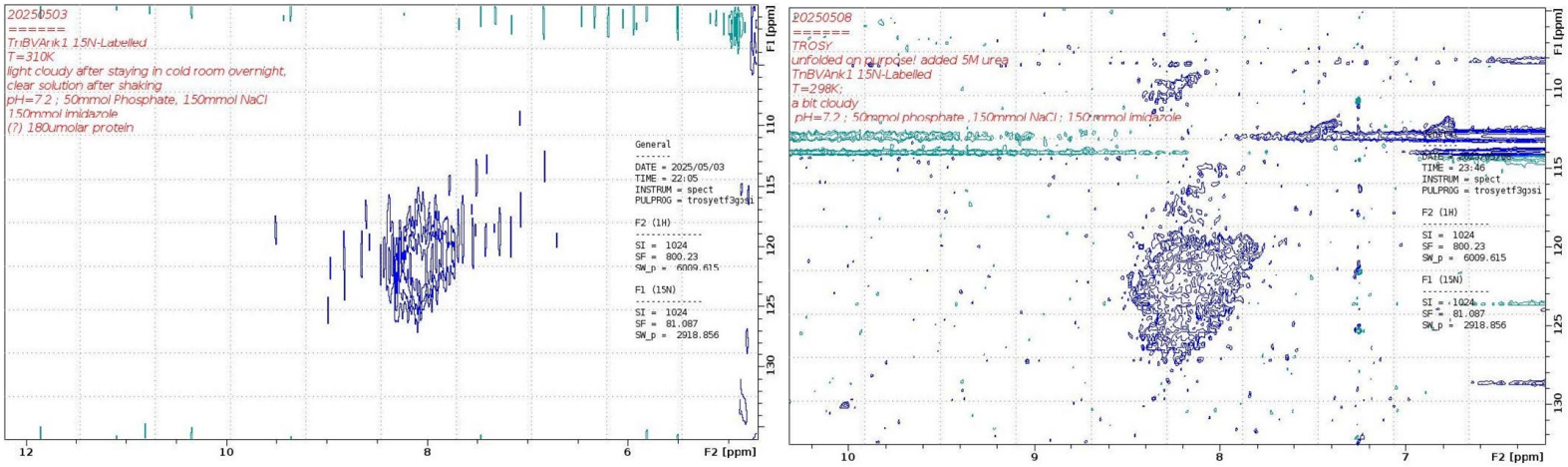

Figure 19: The images show the ^{15}N-^{1}H HSQC spectra results with no 5M UREA added to the sample (A) and with 5M UREA added to the sample (B).

Observing the spectrum without adding 5M UREA we unfortunately observe the presence of a large cluster of strongly overlapping peaks around the central part of the spectrum with only broad peaks of low intensity even after a considerable measurement time. This data is indicative of a significant unstructured and unfolded elements in the protein, and seems to confirm the hypothesis of the presence of some intrinsic disordered regions in the sample.

Therefore, it was decided to obtain further data to confirm this hypothesis by recording a second spectrum ^{15}N-^{1}H HSQC with adding 5M UREA to the sample.

The reason is that if the partial folding hypothesis were true, by adding 5M UREA and denaturing the entire sample, the spectrum should be similar to the spectrum recorded without 5M UREA with a large cluster of strongly overlapping peaks around the central portion of the spectrum. Indeed, as can be seen in Figure 16, the spectrum recorded with 5M UREA also exhibits the same peak arrangement as the spectrum without 5M UREA, providing further evidence supporting the partial folding hypothesis of *Tn*BVANK1.

4. DISCUSSION

The goal of my doctoral project was to perform a functional analysis of *Tn*BVANK1 in an insect model *Drosophila melanogaster*, to understand the potential use of the *Tn*BVANK1 protein as a potential next-generation-environmentally sustainable bioinsecticide. To perform this analysis, *Tn*BVANK1 was ubiquitously expressed in both male and female adult individuals, resulting in decreased viability compared

to controls. Since the phenotype induced by *Tn*BVANK1 expression was more evident in males that ubiquitously express this transgene than in their female counterparts, I performed the analysis in males. A likely explanation for the stronger effect in males is found in the phenomenon of dosage compensation in males individuals of *Drosophila melanogaster*. In flies, dosage compensation leads to a double level of expression of the genes in the X-chromosome in males in comparison to females. In fact, females possess a double dosage of X-genes compared to males, and this mechanism ensures an equal gene expression in the two sexes. This is mediated by the dosage compensation complex (DCC) also known as Male Specific Lethal complex (MSL), due to the male specific lethality phenotype upon loss-of-function of its major components.

At this point, I studied the scientific literature to try to provide a concrete answer to this premature death of adult individuals. I found some articles showing that overexpression of antimicrobial peptides (AMPs) in *Drosophila melanogaster* leads to marked death/shortened adult lifespan. (Badinloo et al., 2018, Kounatidis et al., 2017, and Cao et al., 2013). Specifically, the paper published by (Badinloo et al., 2018) highlights that ubiquitous overexpression of several antimicrobial peptides induces tissue apoptosis, mitochondrial depolarization, and a significant reduction in adult lifespan. This is the most explicit demonstration of AMP toxicity in vivo. Based on these data, it was decided to perform the RT-qPCR analyses of five antimicrobial peptides, in order to observe whether the transcript corresponding to these AMP genes was statistically increased after 14 days of ubiquitous expression of *Tn*BVANK1.

The data show how male flies ubiquitously expressing *Tn*BVANK1 exhibit a significantly accumulation level of all the transcripts at the day 14 of ubiquitous expression of *Tn*BVANK1. These data support the hypothesis of death due to an anomalous overexpression of several antimicrobial peptides.

Observing these surprising qPCR data, we began to hypothesize that in *Drosophila melanogaster*, *Tn*BVANK1 might exert its function differently than in *H. virescens*. The hypothesis is that the ubiquitously overexpression of *Tn*BVANK1 in *Drosophila*, might act as a competitive inhibitor of IkBα by irreversibly binding to NF-κB with its four ankyrin domains.

Interestingly, while previous studies have shown that polydnavirus ANK proteins suppress NF-κB signaling in the host, my data suggest a different outcome in the context of ubiquitous TnBVANK1 expression. The increased antimicrobial peptides expression between day 4 and day 14 days and the reduced adult survival, may indicate a state of immune dysregulation or hyperactivation instead of the canonical suppression. This discordance could reveal functional diversification among ANK proteins. In the natural parasitization of *Heliothis Virescence* by *Toxoneuron nigriceps*, *Tn*BVANK1 expression is spatially and temporally regulated, while ectopic ubiquitous and overexpression in our model *Drosophila melanogaster*, may perturb NF-κB regulatory networks in a way that induce transcriptional activation of immune genes. This preliminary interesting data required further mechanistic studies to clarify whether *Tn*BVANK1 directly modulates NF-κB activity or indirectly influences immune signaling through cellular stress or tissue damage pathways.

Regarding the attempt to resolve the secondary and tertiary structure of *Tn*BVANK1, the results show that the purification of *Tn*BVANK1 was partially successful, as *Tn*BVANK1 co-elutes in protein complexes with ribosomal proteins of both subunits. The buffer solution in which *Tn*BVANK1 elutes after size-exclusion chromatography has pH 7.2. At these pH values, *Tn*BVANK1 has a complete negative surface charge, while the intrinsically disordered regions of the ribosome are positively charged. Therefore, a possible explanation is that during purification by the two chromatographies, *Tn*BVANK1 interacts via electrostatic interactions with these flexible regions of the two ribosomal subunits, preventing complete separation.

Looking at the results obtained by (CD-spectroscopy) it can be seen that the analyzed sample also exhibits 14.9% random coils. This result may be due to both the partial folding of *Tn*BVANK1 and the presence of intrinsically disordered and flexible regions of the proteins present in the ribosomal subunits, which are likely in complex with *Tn*BVANK1.

The ^{15}N - ^{1}H HSQC NMR spectra also appear to support these two hypotheses, since both the spectra with and without 5M UREA exhibit a cluster of overlapping peaks in the center of the spectrum, indicating the presence of random coil regions and unstructured elements in the analyzed sample. The presence of these regions complicates the assignment of resonances and understanding the structure of TnBVANK1.

Overall, my work shows preliminary data that seems to demonstrate that the ubiquitously expression of *Tn*BVANK1 reduces viability of *Drosophila* adult male and female. The lethal phenotype is tightly associated with the overexpression of a number of antimicrobial peptides, which could explain the induced lethality.

The main limitations of this work concern the failure to discover the mechanism through which *Tn*BVANK1 induces AMP activation and adult lethality. A possible hypothesis of the mechanism by which *Tn*BVANK1 induces adult lethality is the hyperactivation of NF-κB. This could continuously activate the transcription of antimicrobial peptides genes, thus inducing the premature death of adult individuals. This needs to be tested at experimental level and will require obtaining a specific antibody to analyze the localization of NF-κB. I expected that NF-κB is constitutively located in the nuclei of TnBVANK1 expressing individuals. Whether my hypothesis is valid, I expected that the knockdown of the *relish* gene, which encodes the *Drosophila* NF-κB gene, should at least in part rescue at AMP activation and the lethal phenotypes. This will require an additional experiment which involves the ubiquitously expression of *Tn*BVANK1 in a *Drosophila* strain with a genetic background that allows NF-κB knockdown.

Furthermore, others preliminary experiments concerning the structural study of *Tn*BVANK1 are required. These data are very important to obtain the protein structure, which is necessary to have the information that are required for modifying the protein sequence and obtain new molecule more active and specific. To get the protein structure, in the future we can first perform ion-exchange chromatography to inhibit the likely electrostatic interaction between *Tn*BVANK1 and the flexible regions of the ribosome. The ultimate goal is

to purify *Tn*BVANK1 and resolve its tertiary structure. Completing the structural study would also allow us to identify any amino acid residues in *Tn*BVANK1 responsible for the interaction with NF-κB.

5. REFERENCES

Casida JE, Quistad GB. Golden age of insecticide research: past, present, or future? Annu Rev Entomol. 1998;43:1-16.

Helgason E, Caugant DA, Olsen I, Kolstø AB. Genetic structure of population of *Bacillus cereus* and *B. thuringiensis* isolates associated with periodontitis and other human infections. J Clin Microbiol. 2000 Apr;38(4):1615-22.

Crickmore N, Zeigler DR, Feitelson J, Schnepf E, Van Rie J, Lereclus D, Baum J, Dean DH. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. Microbiol Mol Biol Rev. 1998 Sep;62(3):807-13.

Wei JZ, Hale K, Carta L, Platzer E, Wong C, Fang SC, Aroian RV. *Bacillus thuringiensis* crystal proteins that target nematodes. Proc Natl Acad Sci U S A. 2003 Mar 4;100(5):2760-5.

Gill SS, Cowles EA, Pietrantonio PV. The mode of action of *Bacillus thuringiensis* endotoxins. Annu Rev Entomol. 1992;37:615-36.

Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, Feitelson J, Zeigler DR, Dean DH. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol Mol Biol Rev. 1998 Sep;62(3):775-806.

Aronson AI, Shai Y. Why *Bacillus thuringiensis* insecticidal toxins are so effective: unique features of their mode of action. FEMS Microbiol Lett. 2001 Feb 5;195(1):1-8.

de Maagd RA, Bravo A, Crickmore N. How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. Trends Genet. 2001 Apr;17(4):193-9.

CHOMA, C.T., SUREWICZ, W.K., CAREY, P.R., POZSGAY, M., RAYNOR, T. and KAPLAN, H. (1990), Unusual proteolysis of the protoxin and toxin from *Bacillus thuringiensis*. European Journal of Biochemistry, 189: 523-527.

Ferré J, Van Rie J. Biochemistry and genetics of insect resistance to *Bacillus thuringiensis*. Annu Rev Entomol. 2002;47:501-33.

Dahlman, D.H. and Vinson, S.B. (1993) Teratocytes: Developmental and Biochemical Characteristics. In: Beckage, N.E., Thompson, S.N. and Federici, B.A., Eds., Parasites and Pathogens of Insects, Vol. 1, Academic Press, New York, 145-166.

Francesco Pennacchio, S. Bradleigh Vinson, Ermenegildo Tremblay, Angela Ostuni, Alteration of ecdysone metabolism in *Heliothis virescens* (F.) (Lepidoptera: Noctuidae) larvae induced by *Cardiochiles nigriceps* Viereck (Hymenoptera: Braconidae) teratocytes, Insect Biochemistry and Molecular Biology, Volume 24, Issue 4, 1994, Pages 383-394, ISSN 0965-1748.

T. Tanaka, S. B. Vinson, Interaction of Venoms with the Calyx Fluids of Three Parasitoids, *Cardiochiles nigriceps*, *Microplitis croceipes* (Hymenoptera: Braconidae), and *Campoletis sonorensis* (Hymenoptera: Ichneumonidae) in Effecting a Delay in the Pupation of *Heliothis virescens* (Lepidoptera: Noctuidae), Annals of the Entomological Society of America, Volume 84, Issue 1, 1 January 1991, Pages 87–92.

Francesco Pennacchio, Patrizia Falabella, Rocco Sordetti, Paola Varricchio, Carla Malva, S Bradleigh Vinson, Prothoracic gland inactivation in *Heliothis virescens* (F.) (Lepidoptera:Noctuidae) larvae parasitized by *Cardiochiles nigriceps* Viereck (Hymenoptera:Braconidae), Journal of Insect Physiology, Volume 44, Issue 9, 1998, Pages 845-857, ISSN 0022-1910.

Paola Varricchio, Patrizia Falabella, Rocco Sordetti, Franco Graziani, Carla Malva, Francesco Pennacchio, *Cardiochiles nigriceps* polydnavirus: molecular characterization and gene expression in parasitized *Heliothis virescens* larvae, Insect Biochemistry and Molecular Biology, Volume 29, Issue 12, 1999, Pages 1087-1096, ISSN 0965-1748.

Falabella P, Varricchio P, Provost B, Espagne E, Ferrarese R, Grimaldi A, de Eguileor M, Fimiani G, Ursini MV, Malva C, Drezen JM, Pennacchio F. Characterization of the IkB-like gene family in polydnaviruses associated with wasps belonging to different Braconid subfamilies. J Gen Virol. 2007 Jan;88(Pt 1):92-104.

Provost B, Varricchio P, Arana E, Espagne EFalabella P, Huguet E, La Scaleia R, Cattolico L, Poirié M, Malva C, Olszewski JA, Pennacchio F, Drezen J 2004. Bracoviruses Contain a Large Multigene Family Coding for Protein Tyrosine Phosphatases. J Virol 78:.

Falabella, P., Caccialupi, P., Varricchio, P., Malva, C. and Pennacchio, F. (2006), Protein tyrosine phosphatases of *Toxoneuron nigriceps* bracovirus as potential disrupters of host prothoracic gland function †. Arch. Insect Biochem. Physiol., 61: 157-169.

Ghosh S, May MJ, Kopp EB. NF-kB and Rel proteins: evolutionarily conserved mediators of immune responses. Annu Rev Immunol. 1998;16:225-60.

Ghosh S, Karin M. Missing pieces in the NF-kB puzzle. Cell. 2002 Apr;109 Suppl:S81-96.

Webb BA, Strand MR, Dickey SE, Beck MH, Hilgarth RS, Barney WE, Kadash K, Kroemer JA, Lindstrom KG, Rattanadechakul W, Shelby KS, Thoetkiattikul H, Turnbull MW, Witherell RA. Polydnavirus genomes reflect their dual roles as mutualists and pathogens. Virology. 2006 Mar 30;347(1):160-74.

Rogers S, Wells R, Rechsteiner M. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science. 1986 Oct 17;234(4774):364-8.

Kumar A, Balbach J. Folding and Stability of Ankyrin Repeats Control Biological Protein Function. Biomolecules. 2021 Jun 5;11(6):840.

Kent S Shelby, Bruce A Webb, Polydnavirus-mediated suppression of insect immunity, Journal of Insect Physiology, Volume 45, Issue 5, 1999, Pages 507-514, ISSN 0022-1910.

Kent S. Shelby, Olushola A. Adeyeye, B.Moses Okot-Kotber, Bruce A. Webb, Parasitism-Linked Block of Host Plasma Melanization, Journal of Invertebrate Pathology, Volume 75, Issue 3, 2000, Pages 218-225, ISSN 0022-2011.

Di Lelio I, Caccia S, Coppola M, Buonanno M, Di Prisco G, et al. (2014) A Virulence Factor Encoded by a Polydnavirus Confers Tolerance to Transgenic Tobacco Plants against Lepidopteran Larvae, by Impairing Nutrient Absorption. PLOS ONE 9(12): e113988.

Bitra K, Suderman RJ, Strand MR (2012) Polydnavirus Ank Proteins Bind NF-κB Homodimers and Inhibit Processing of Relish. PLOS Pathogens 8(5): e1002722.

Salvia, R., Grossi, G., Amoresano, A. et al. The multifunctional polydnavirus *Tn*BVANK1 protein: impact on host apoptotic pathway. Sci Rep 7, 11775 (2017).

Missotten M, Nichols A, Rieger K, Sadoul R. Alix, a novel mouse protein undergoing calcium-dependent interaction with the apoptosis-linked-gene 2 (ALG-2) protein. Cell Death Differ. 1999 Feb;6(2):124-9.

The mechanism utilized by *Toxoneuron nigriceps* in inhibiting the host immune system, Patrizia Falabella, ISJ-Invertebrate Survival Journal, 2018, volume 15, Pages 240-255.

Duchi S, Cavaliere V, Fagnocchi L, Grimaldi MR, Falabella P, Graziani F, Gigliotti S, Pennacchio F, Gargiulo G. The impact on microtubule network of a bracovirus IkB-like protein. Cell Mol Life Sci. 2010 May;67(10):1699-712.

Valzania L, Romani P, Tian L, Li S, Cavaliere V, et al. (2014) A Polydnavirus ANK Protein Acts as Virulence Factor by Disrupting the Function of Prothoracic Gland Steroidogenic Cells. PLOS ONE 9(4): e95104.

McGuire SE, Le PT, Osborn AJ, Matsumoto K, Davis RL. Spatiotemporal rescue of memory dysfunction in *Drosophila.* Science. 2003 Dec 5;302(5651):1765-8.

Badinloo M, Nguyen E, Suh W, Alzahrani F, Castellanos J, Klichko VI, Orr WC, Radyuk SN. Overexpression of antimicrobial peptides contributes to aging through cytotoxic effects in *Drosophila* tissues. Arch Insect Biochem Physiol. 2018 Aug;98(4):e21464.

Kounatidis I, Chtarbanova S, Cao Y, Hayne M, Jayanth D, Ganetzky B, Ligoxygakis P. NF-κB Immunity in the Brain Determines Fly Lifespan in Healthy Aging and Age-Related Neurodegeneration. Cell Rep. 2017 Apr 25;19(4):836-848.

Y. Cao,S. Chtarbanova,A.J. Petersen, & B. Ganetzky, Dnr1 mutations cause neurodegeneration in *Drosophila* by activating the innate immune response in the brain, Proc. Natl. Acad. Sci. U.S.A. 110 (19) E1752-E1760.

Wimberly BT, White SW, Ramakrishnan V. The structure of ribosomal protein S7 at 1.9 A resolution reveals a beta-hairpin motif that binds double-stranded nucleic acids. Structure. 1997 Sep 15;5(9):1187-98.

Sergey Melnikov, Kasidet Manakongtreecheep, Dieter Söll, Revising the Structural Diversity of Ribosomal Proteins Across the Three Domains of Life, Molecular Biology and Evolution, Volume 35, Issue 7, July 2018, Pages 1588–1598.

Timsit Y, Acosta Z, Allemand F, Chiaruttini C, Springer M. The role of disordered ribosomal protein extensions in the early steps of eubacterial 50 S ribosomal subunit assembly. Int J Mol Sci. 2009 Mar;10(3):817-834.

Peng Z, Oldfield CJ, Xue B, Mizianty MJ, Dunker AK, Kurgan L, Uversky VN. A creature with a hundred waggly tails: intrinsically disordered proteins in the ribosome. Cell Mol Life Sci. 2014 Apr;71(8):1477-504.

Pennacchio F, Strand MR. Evolution of developmental strategies in parasitic hymenoptera. Annu Rev Entomol. 2006;51:233-58.