THE MOLECULAR CHARACTERIZATION OF *TRICHOPLUSIA NI* SINGLE NUCLEOCAPSID NUCLEOPOLYHEDROVIRUS: A STUDY ON EARLY REGULATORY FEATURES

BY

WEIZHOU WANG

A thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Department of Biotechnology University of the Western Cape

UNIVERSITY of the
WESTERN CAPE

Supervisor: Professor S. Davison

July 2003

https://etd.uwc.ac.za/

THE MOLECULAR CHARACTERIZATIN OF *TRICHOPLUSIA NI* SINGLE NUCLEOCAPSID NUCLEOPOLYHEDROVIRUS: A STUDY ON EARLY REGULATORY FEATURES

BY

WEIZHOU WANG

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biotechnology University of the Western Cape

UNIVERSITY of the
WESTERN CAPE

Supervisor: Prof. Sean Davison

July 2003

https://etd.uwc.ac.za/

DECLARATION

I declare that the research work titled "The molecular characterization of *Trichoplusia ni* single nucleocapsid nucleopolyhedrovirus: a study on early regulatory features" is my own work and all the resources I have used or quoted have been indicated and acknowledged by means of complete references.

Full name ___Weizhou Wang___

Signed ____________ Date ___July 24, 2003______

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Acknowledgements

I would like to express my sincere gratitude and appreciation to the following:

Prof. Sean Davison, my supervisor, whose inspiration, encouragement, guidance and great help, not only through this project but also in my personal life, will always be remembered.

Prof. Chris Gehring for providing me access to his laboratory and for his encouragement and help.

Prof. Peter Krell for allowing me to work in his laboratory for a period at the Department of Microbiology, University of Guelph, Canada

Colleagues in the virology laboratory for the valuable discussions, advice and assistance towards part of the project.

Steve Lord and J. D. de Jong, Department of Microbiology, University of Guelph, Canada, for technical assistance towards part of the work.

Pastor Arthur Volker and Mrs. Joy Volker, Mowbray Baptist Church, for their love, spiritual guidance and invaluable help to myself and my family which have strengthened and encouraged me to face difficult challenges during my studies.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

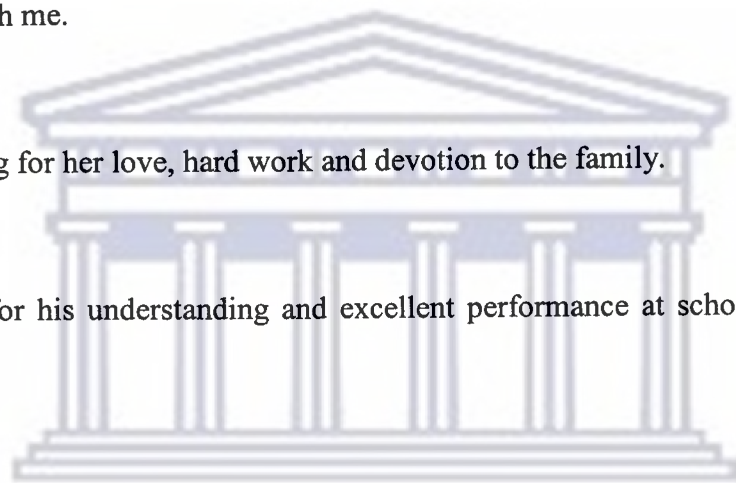
Mrs. Gill Lanham and Mr. John Lanham for their kind caring for my family, especially for their love towards my son.

Mr. Don Waterson and Mrs. Sheila Waterson, Mowbray Baptist Church, for their special care and friendliness towards my family and myself.

Friends at the Mowbray Baptist Church, for their encouragement, fellowship and sharing God's Words with me.

My wife Yanqing for her love, hard work and devotion to the family.

My son Zijing for his understanding and excellent performance at school for which we are always proud.



https://etd.uwc.ac.za/

List of Figures

Figure 1.1 Structural diagram of the two baculovirus pheontypes........................4

Figure 1.2 Schematic diagram of the baculovirus life cycle...............................8

Figure 2.1 The location of the *ie-1* gene within an 11 kb *Eco*RI fragment and an adjoining 2.3 kb *Eco*RI fragment...33

Figure 2.2 Nucleotide and deduced amino acid sequences of the TnSNPV *ie-1* gene...36

Figure 2.3 Multiple sequence alignment of *ie-1* promoter regions.......................38

Figure 2.4 Alignment of amino acid sequences of eleven baculovirus *ie-1* genes......42

Figure 2.5 Acid-base profiles of the eleven baculovirus IE-1 proteins..................44

Figure 2.6 Phylogenetic relationships of the predicted full length (a) and the C-terminal two thirds (b) of the IE-1 proteins..................................48

Figure 3.1 Schematic representation and amino acids sequences of the wild type IE-1s and their truncations...64

Figure 3.2 Structural alignments of the putative C-terminal binding regions and the b-HLH-like domains of eight IE-1s.......................................70

Figure 3.3 Expression, purification and detection of the AcNPV and TnSNPV IE-1 truncations...71

Figure 3.4 Efficacy of purification of the 6xHis tagged truncation (Δ1-452)..........72

Figure 3.5 DNA-binding activities of the IE-1 truncation and controls.................74

Figure 4.1 The location of the me-53 gene within an 11 kb and an adjoining 2.3 kb *Eco*RI fragment and the AcMNPV recombinant..........................86

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Figure 4.2 Nucleotide and deduced amino acid sequences of the TnSNPV *me-53*-containing fragment..92

Figure 4.3 Alignment of the deduced amino acid sequences of twelve baculovirus *me-53* genes..97

Figure 4.4 Pattern of the zinc finger-like motifs in the baculovirus ME-53s.............99

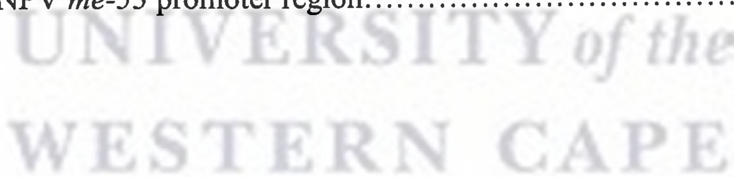
Figure 4.5 Uninfected and AcBacTnme53GFP-infected Sf 9 and High5 cells.........100

Figure 4.6 Temporal expression of the TnSNPV *me-53* transcripts by RT-PCR analysis..103

Figure 4.7 Comparison of regulatory elements within the upstream regions of ten NPV *me-53s*..105

Figure 4.8 Rooted neighbor-joining trees of deduced baculovirus ME-53 (a) and IE-1 (b) proteins..108

Figure 4.9 Point mutation of the TATA motifs and the RNA start site of the TnSNPV *me-53* promoter region...116



https://etd.uwc.ac.za/

List of Tables

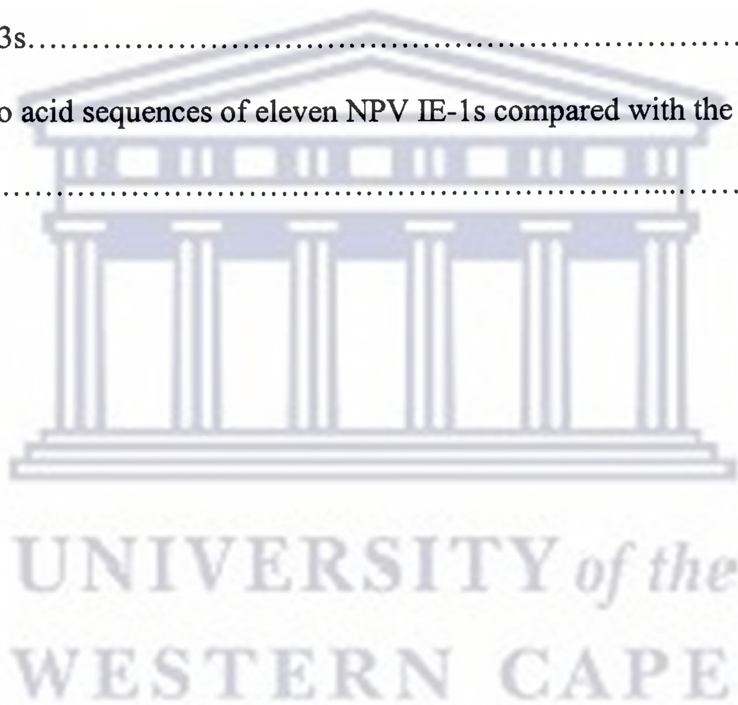
Table 2.1 The N-terminal net negative charge regions of eleven baculovirus IE-1s.....45

Table 2.2 Identity values of the amino acid sequences of eleven baculovirus IE-1s.....46

Table 2.3 Comparison of full length (a) and the C-terminal two third portions (b) of ten deduced NPV IE-1s with the TnSNPV homologue.......................47

Table 4.1 Identity values of the amino acid sequences of twelve baculovirus ME-53s..106

Table 4.2 Amino acid sequences of eleven NPV IE-1s compared with the TnSNPV IE-1...106



https://etd.uwc.ac.za/

CONTENTS

Declaration.....I

Acknowledgements.....II

List of Figures.....IV

List of Tables.....VI

Chapter 1 General Introduction.....1

1.1 Introduction to Baculovirus.....2

1.2 Taxonomy and Structure.....3

1.3 Infection Cycle.....6

1.4 Genome Organization.....9

1.5 Gene Regulation, Expression and DNA Replication.....12

1.6 Genetic Engineering and Expression Vectors of Baculoviruses.....16

1.7 *Trichoplusia ni* and TnSNPV.....18

1.8 Scope of the Thesis.....19

Chapter 2 Identification, Sequence Analysis and Phylogeny of the Immediate Early Gene 1 of the *Trichoplusia ni* Single Nucleocapsid Polyhedrosis Virus.....25

2.1 Abstract.....26

2.2 Introduction.....27

2.3 Materials and Methods.....31

2.3.1 Insects, virus isolation and purification.....31

2.3.2 DNA extraction, manipulation and suquencing.....32

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

2.3.3 Computer analysis..33

2.4 Results..34

2.4.1 Nucleotide and protein sequence analyses..34

2.4.2 Comparison and phylogenetic analysis of IE-1 proteins..46

2.5 Discussion..49

Chapter 3 Experimental Prediction of a DNA-Protein Binding Region of the *Trichoplusia ni* Single Nucleocapsid Nucleopolyhedrvirus IE-1.......58

3.1 Abstract..59

3.2 Introduction..60

3.3 Materials and Methods..62

3.3.1 Prediction of the C-terminal secondary structures of NPV IE-1s.......62

3.3.2 Plasmid constructs..63

3.3.3 Protein expression and purification..65

3.3.4 Western blot assay..66

3.3.5 Radio-labeling of the probe and electrophoretic mobility shift assay..67

3.4 Results..68

3.4.1 Prediction of the secondary structures and the experimental design....68

3.4.2 Expression, purification and detection of the AcNPV and TnSNPV IE-1 truncations..71

3.4.3 Gel shift assay..73

3.5 Discussion..75

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Chapter 4 Identification and Characterization of a Major Early-Transcribed Gene of *Trichoplusia ni* Single Capsid Nucleopolyhedrovirus Using the Baculovirus Expression System......80

4.1 Abstract......81

4.2 Introduction......82

4.3 Materials and Methods......85

4.3.1 DNA manipulation of the TnSNPV *me-53*, sequencing and Computer analysis......85

4.3.2 Cells and virus......87

4.3.3 Generation of the recombinant baculoviruses......88

4.3.4 Total RNA isolation, 5' and 3' RACE and RT-PCR analyses of the TnSNPV *me-53* transcripts......89

4.3.5 Phase contrast and fluorescence microscopy......94

4.4 Results......94

4.4.1 Sequence analysis of the TnSNPV *me-53*......94

4.4.2 Infection of the recombinant viruses and GFP expression......99

4.4.3 Transcription analyses of the TnSNPV and authentic *me-53s* of AcBacTnme53GFP......101

4.4.4 Analysis of the 5' non-coding regions of the NPV *me-53s*......103

4.4.5 Comparative phylogenetic analysis of IE-1 and ME-53 proteins......105

4.5 Disscusion......108

4.5.1 Amino acid sequence analysis of the TnSNPV ME-53......108

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

4.5.2 Accurate transcription initiation of the TnSNPV *me-53* Promoter....110
4.5.3 Early transcriptional analysis of the TnSNPV *me-53*....111
4.5.4 Early regulatory elements of the baculovirus *me-53*s and strategy for determination of the functional role of the TATA motifs in the TnSNPV *me-53* promoter region....113
4.5.5 Late transcription analysis of the TnSNPV *me-53*....118
4.5.6 Late transcription initiation element of the NPV *me-53s*....119
4.5.7 Host factor binding sites of the NPV *me-53s*....120
4.5.8 Phylogenetic analysis of the TnSNPV ME-53....121
4.5.9 A potential model system for improving bio-insecticidal agents...123
Chapter 5 Summary and Concluding Remarks....125
References....132
List of Publications....171

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Chapter 1

UNIVERSITY of the
WESTERN CAPE

General Introduction

https://etd.uwc.ac.za/

1.1 Introduction to Baculoviruses

Baculoviridae are a diverse family of insect virus pathogens. They are infectious for arthropods but the majority of the infecting hosts are insects of the order Lepidoptera (Blissard, 1996). Some baculoviruses have also been isolated from the other insect orders Hymenoptera, Diptera and Coleoptera, as well as from the crustacean order Decapoda (Rohrmann, 1992). Although over 800 species of insects are infected by baculoviruses, individual isolates normally show a limited host range and infect only closely related insects within a single order (Ijkel, 2001). Due to the characteristic of causing epizootic in nature (the viruses infect insect hosts as an epidemic does to humans), which may play a role in controlling insect population, baculoviruses have attracted more and more attention in scientific research and become more popular as biological control agents. Unlike wide-spectrum chemical insecticides with disadvantages such as non-specific insect targets, causing chemical resistance in insects and environmental damage, baculoviruses have provided an ecological advantage as an alternative to chemical insecticides with little direct effect on human health, hazard to other vertebrates and non-resistant to insects (Persley, 1996).

Although baculoviruses have potential as biological insecticides, they also have limitation in insect control. Many attempts have been made to improve molecular properties of baculoviruses as biological pesticides for pest control. The technologies for manipulation and genetic engineering of the viral genome have been successfully used to improve molecular properties of baculoviruses as biological pesticides and also allow to

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

develop baculovirus expression vectors which have been widely used for production of recombinant protein of pharmaceutical interest as well as for basic and applied biomedical and veterinary research purposes (Possee, 1997; Kost and Condreay, 1999).

To date, only some of baculoviruses have been identified. The most intensively studied baculovirus is *Autographa californica* nucleocapsid polyhedrosis virus (AcMNPV). A wealth of information is available for this virus and other identified baculoviruses. In recent years a South African *Trichoplusia ni* single capsid nuclear polyhedrosis virus (TnSNPV) isolate has been partially characterized, but not much is known about the virus, especially for its gene regulation. This chapter will provide some background information concerning the biology and molecular genetics of baculoviruses in general and TnSNPV in particular.

1.2 Taxonomy and Structure

The family Baculoviridae is divided into two taxonomic genera: Nucleopolyhedrovirus (NPV) and Granulovirus (GV) (Murphy *et al.*, 1995), which is based on the number of virions occluded in a protein matrix called the occlusion body, polyhedron or granulum. In NPVs several or as many as 200 virions are found in one polyhedron. In GV a single or at most a few virions are occluded within a granulum. Polyhedron and granulum mainly consist of the two proteins, respectively, named polyhedrin and granulin. Polyhedrin and granulin are structurally and functionally related proteins. NPVs have two morphotypes, one referred to as single enveloped nucleocapsid NPV (SNPV) and other multiple enveloped nucleocapsid NPV (MNPV). In the latter, up to nine nucleocapsids

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

are assembled in a single virion envelope before several of these packages are occluded into a polyhedron (van Strien, 1997). In contrast, GVs contain a single nucleocapsid per envelope. The other feature of NPVs is that they have two virion phenotypes. The virions released from the occlusion bodies are termed occlusion-derived virus (ODV), while others produced in epithelial cells of the midgut and found in the tracheal system and hemocoel of infected insect host is termed budded virus (BV). The structural difference between ODV and BV is shown in Fig. 1.1.

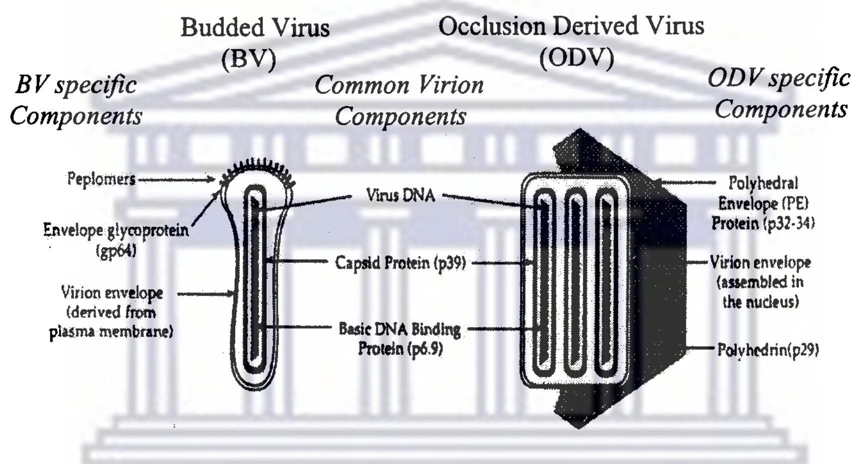


Figure 1.1 Structural diagram of the two baculovirus phenotypes (Blissard and Rohrmann, 1990). The locations of relevant virion components are indicated by arrows and solid lines.

UNIVERSITY of the
WESTERN CAPE

To date, twelve different baculovirus species have been classified into NPV genus and five to GV genus (Blissard *et al*., 2000). All the baculoviruses are named after the host from which they were first isolated. To establish the identity of baculovirus species, individual genes such as polyhedrin/granulin and DNA polymerase (Zanotto *et al*., 1993; Cowan *et al*., 1994; Bluach *et al*., 1999; Bideshi *et al*., 2000) have been used to study genetic relatedness of baculoviruses in the recent years. The phylogenetic analyses of

https://etd.uwc.ac.za/

several other genes also contributed to infer baculovirus phylogenies. These analyses proposed that baculoviruses were separated into two groups, NPVs and GVs, while NPVs were further subdivided into group I and II. However, phylogenies based on different genes are often in conflicts with one another (Herniou *et al.*, 2001). With availability of increased complete genome sequence data, the different approaches such as genome sequence comparison, gene content and gene order have been used to infer baculovirus and other virus phylogenies (Hannenhalli et al., 1995; Hu, Z. H., 1998; Montague and Hutchison, 2000; Herniou at al., 2001, 2003). Herniou *et al* (2001) combined three analyses based on trees constructed from data sets of gene order, gene content and a large number of individual genes shared by nine baculoviruses, to study baculovirus phylogenies. Although the latter has been shown to be the most robust approach for inference of baculovirus phylogenies, as suggested by Herniou *et al* (2001), two other different data sets that are phylogenetically informative should be used to support the combined gene analysis. More recently, reconstruction of baculovirus phylogenies was conducted by Herniou *et al* (2003). The data set generated from a core set of 30 genes conserved among thirteen baculovirus genomes, including the dipteran NPV (*Culex nigripalpus* NPV: CuniNPV), was used to yield a majority rule consensus tree with four major groups, the GV, group I and II lepidopteran baculovirus NPV and the dipteran NPV (CuniNPV). This phylogeny was supported by analyses of gene order and gene content. The comparative genomic approaches have provided correct resolution to baculovirus phylogenies and will give insights into the understanding of baculovirus biology, host-interaction and evolution (Herniou *et al* 2003).

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

1.3 Infection Cycle

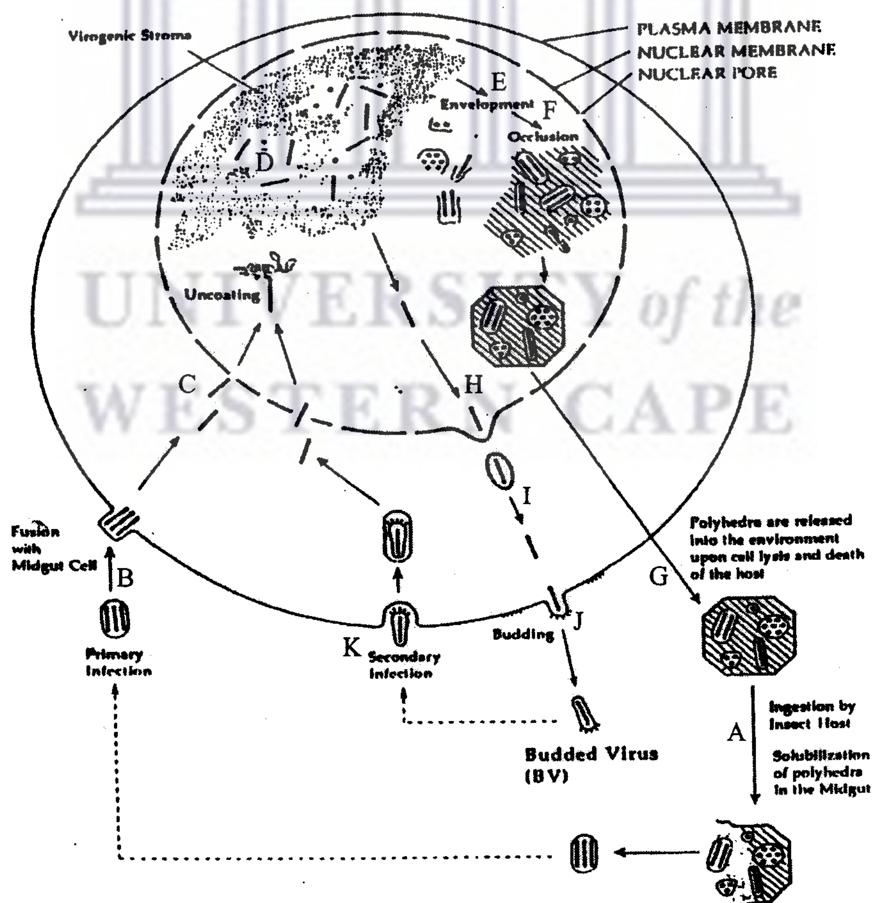
In the baculovirus life cycle, two structurally distinct phenotypes play specific roles in the infection process (Fig. 1.2). Transmission of baculovirus in an insect population occurs via either ingestion of food or other sources (soil, water, or crevices of plants) contaminated with occlusion bodies (OBs) in the environment (Blissard *et al.*, 2000). During the ingestion the OBs rapidly dissolve in the alkaline environment of the larval midgut and then released enveloped virions termed 'occlusion-derived virus' (ODV) enter the host cells by fusion of the virion envelope with microvilli of midgut columnar epithelial cells (Granados and Lawler, 1981; Horton and Burand, 1993). After fusion process, the nucleocapsids (NCs) are released into the cytoplasm, transported into, and uncoated in the nucleus where viral gene expression, DNA replication and assembly of progeny NCs occur. The newly assembled progeny nucleocapsids then bud through the nuclear membrane and subsequently through plasma membrane on the basal side of the epithelial cells into hemocoel or tracheoblast, becoming budded virions (BVs). During the budding process, nucleocapsids become enveloped by the plasma membrane where a glycoprotein, *gp64*, is incorporated (Volkman *et al.*, 1984; Blissard and Rohrmann, 1989; Oomens *et al.*, 1995). This protein is required for efficient budding (Monsma *et al.*, 1996; Oomens and Blissard, 1999). When BVs spread through tracheal systems and hemolymph to other tissues (fat body, muscle, salivary gland, gonads and cuticular epidemi, etc.) and enter other cells by adsorptive endocytosis, a secondary infection and viral DNA replication occur. The larvae insect tracheal system was found to be mainly responsible for virus systemic infection (Engelhard *et al.,* 1994). Its lymph system and

https://etd.uwc.ac.za/

epidermal feet (cytoplasmic extensions, which is interdigitated to form the contact among tracheal epidermis) are likely to facilitate the rapid systemic spread of virus because the lymph channels and the epidermal feet could facilitate tissue-specific cell communication over long distances (Locke, 1985), while tracheoblasts (newly produced daughter cells of the tracheal epidermis on tracheoles) facilitate immediate secondary infection by crossing the basal lamina barriers and serve as principle carriers to deliver virus to the insect tracheal epidermal cells and to hemocoel, because tracheoblasts directly contact the tracheal epidermis and are also closely associated with other tissues. BVs are essential for systemic infection of insect hosts (tissue to tissue and cell to cell) (Keddie *et al.*, 1989; Granados and Lawler, 1981), while the larvae insect tracheal system is mainly involved in the process. It provides not only a productive target tissue but also a conduit for facilitating movement of the virus from midgut to other tissues (Engelhard *et al.,* 1994). The ODVs, the virions formed later in the infected cells, are not released by budding. They are enveloped *de novo* in the nucleus and later occluded with polyhedrin (polyhedra-derived virus, PDV). Typically, occlusion bodies are first observed at 24 h p.i. (Blissard and Rohrmann, 1990). Except for being responsible for the initiation of an infection in the insect, the ODV provides protection for the virus against harsh environment and can allow the virus to remain viable outside of the hosts for years. Although the two virion phenotypes (BV and ODV) appear to be genetically identical (Smith and Summer, 1978; Blissard, 1996), they differ in morphology, protein composition, source of virion envelopes (Fig. 1.1), tissue specificity and mode of viral entry into host cells.

https://etd.uwc.ac.za/

It is not until the final stage of infection that infected larvae become sluggish and cease feeding. The disintegration of the larvae occurs and is enhanced by two virus-encoded enzymes, chitinase and cathepsin. Both of them are associated with liquefaction of the host at the end of the infection process. The insect cuticle, the structure of which is formed through chitin crosslinked with sclerotized proteins, must be degraded in this process. Cathepsin may trip the protein from cuticular chitin so that chitinase can act on the naked chitin and then cuticle degrades (Ohkawa, et al., 1994; Hawtin, *et al*, 1995; Hawtin, *et al*, 1997; Slack *et al.*, 1995; Hill, *et al.*, 1995). After cells and tissues disintegrate, OBs are released into the environment and the next infection cycle begins as occlusion bodies are ingested by other susceptible insects.



https://etd.uwc.ac.za/

Figure 1.2 Schematic diagram of the baculovirus life cycle (from van der Beek, 1980). Ingested polyhedra (OB) are solubilized in the midgut and virions are released (A). Virions of the occulusion-derived virus (ODV) phenotype are released and enter midgut epithelial cells by fusion with microvilli (B). Nucleocapsids are transported to the nucleus and uncoated, followed by gene expression and viral replication (C). Progeny nucleocapsids are synthesized and assembled in the virogenic stroma (D). Following envelopment in the nucleus (E), progeny nuclecapsids go through the budding process (H, I, J). Budded virions infect other host cells by endocytosis (K). Nucleocapsids produced in the late stages of an infection are occluded within polyhedrin protein (F) and become matured form. Upon insect death and cell lysis, occlusion bodies are released into the environment (G).

1.4 Genome Organization

UNIVERSITY of the
WESTERN CAPE

Baculoviruses have circular, supercoiled, double-stranded DNA genomes that vary in size from 90 to180 kb depending on the species. Until now, genome sequences of ten lepidopteran NPVs, a dipteran NPV and four GVs are available. They are: *Autographa califonica* NPV (NC001623) (Ayres *et al.*, 1994); *Bomyx mori* NPV (NC001962) (Gomi *et al.*, 2001); *Orgyia pseudotsugata* NPV (NC001875) (Ahrens *et al.*, 1997); *Helicoverpa armigera* NPV (NC003094) (Chen *et al.*, 2001); *Helicoverpa zea* NPV (AF334030) (Chen *et al.*, 2002); *Spodoptera litura* NPV (NC003102) (Pang *et al* 2001); *Spodoptera exigua* NPV (NC002169) (Ijkel *et al.*, 1999); *Lymantria dispar* NPV (NC001973) (Kuzio

https://etd.uwc.ac.za/

et al., 1999); *Mamestra configurata* NPV A (NC003529) (Erlandson *et al.*, 2002); *Epiphyas postvittana* NPV (NC003083) (Hyink *et al.*, 2001); *Culex nigripalpus* NPV (AF403738) (Afonso *et al.*, 2001); *Xestia c-nigrum* GV (NC002331) (Hayakawa *et al.*, 1999); *Plutella xylostella* GV (NC002593) (Hashimoto *et al.*, 2000); *Phthorimaea operculella* GV (NC004062) (Croizier *et al.*, unpublished); *Cydia pomonella* GV (NC001623) (Luque *et al.*, 2001); The sizes of these genome range from 101(PxGV) to 179 (XcGV) kb. The number of potential open reading frames (ORF) differs from 109 (CuniNPV) to 181(XcGV) in the different genomes.

Although genome arrangements among either group I NPVs (AcMNPV, BmNPV and OpMNPV) or GVs are similar (Blissard and Rohrmann, 1990; Hu et al., 1998; Herniou et al., 2003), gene order, gene content and gene conservation among different baculovirus groups have shown a high degree of variability. The comparison of genomic sequences of either 9 or 13 baculoviruses (Herniou et al. 2001; 2003) revealed that gene order is quite different among different baculovirus groups, and even among group II NPVs, while gene content showed that 17 genes are only found in group I NPVs, 14 genes specific to lepidopteran NPVs and 27 to GVs. These studies indicated that 63 genes are conserved among GVs, group I and II NPVs, while the dipteran NPV (CuniNPV) only shares 30 genes with three other group baculoviruses. The differences of genomes in different data sets reflect differences of both biological features and evolutionary relatedness among baculovirus groups (Ijkel, 2001; Herniou et al., 2003). During the course of baculoviruse evolution, some genes have been lost and some acquired, and extensive genome rearrangements have occurred. Although these changes occurred during the baculovirus

https://etd.uwc.ac.za/

evolution, the cluster of four genes were preserved in the same relative position in all the sequenced genomes. Two of them (*helicase* and *lef-5*) are essential for baculovirus replication and transcription, and the functions of two others are unknown. The conservation of the gene order suggested that some essential functional requirements in the virus infection cycle may be required to preserve the gene arrangement (Herniou et al., 2003). With the growth of knowledge in baculovirus gene function, differences of gene conservation and genome rearrangements among baculovirueses will be better understood.

Among baculoviruses the AcMNPV is most intensively studied. It has a genome size of 133,894 bp, encoding 154 proteins (Ayres *et al*., 1994). Several homologous regions (*hrs*) with repeat sequences are interspersed at specific sites along the length of the genome. Hrs are believed to be origins of DNA replication and serve as enhancers of gene expression. They are well conserved among all the baculoviruses, including CuniNPV (Afonso, et al., 2001), but excluding CpGV which does not have the homologues of baculovirus typical *hrs*. However, several different classes of repeats were found in the CpGV genome, the functions of which have not been identified (Luque et al., 2001). Although the number of *hrs* varies among different baculovirus species, ranging from 4 (PxGV) to 13 (LdMNPV) (Hayakawa et al., 2000), with the exception of CpGV, they are generally presented in baculovirus genomes, especially in NPVs, and therefore, can be considered as a characteristic feature of baculoviruses (Chen, 2001; Ijkel, 2001).

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

1.5 Gene Regulation, Expression and DNA Replication

Baculovirus gene expression is temporally regulated in a cascaded manner in which each successive phase is dependant on the previous one (Blissard and Rohrmann, 1990). Three main classes of genes are recognized: early, late and very late (Blissard and Rohrmann, 1990; 1991). Early genes are transcribed by host RNA polymerase II, while late genes by a viral encoded RNA polymerase (Guarino *et al.*, 1998). The typical eukaryotic consensus transcription motifs, such as a TATA box and a CATG motif, usually function as core elements of baculovirus early promoters. TATA motif plays a critical role in regulating the rate of transcription and selecting the RNA start site for early gene transcription, especially in the absence of CAGT motif (Dickson and Friesen, 1991; Guarino and Smith, 1992; Blissard et al., 1992), while CAGT motif not only serves as early transcription start site for many early genes but also as an initiator in the case of the AcMNPV *ie-1* (Pullen and Friesen, 1995). Unlike TATA and CAGT motifs responsive to host factors in uninfected cells, an unusual promoter motif, CGTGC, positioned at the RNA start site of *dnapol* or *p143,* is inactive in uninfected cells but highly responsive to transregulator IE-1 for *dnapol* (Ohresser et al., 1995) and to IE-1, IE-2, PE38 and *cis*-liked *hr5* enhancer for *p143* (Lu and Carstens, 1993). In addition to the early promoter motifs, two *cis*-acting elements, CGT and GC, presented in upstream activating regions (UAR) of *p35, 39k* and *gp64,* stimulate basal promoter activities from these genes (Dickson and Friesen, 1991; Guarino and Smith, 1992; Kogan and Blissard, 1994). Another *cis*-acting motif in UAR is GATA that was identified as a nuclear factor binding site in host insect cells. Mutation analysis of the motif resulted in reduction of the

https://etd.uwc.ac.za/

promoter activity in the case of *gp64,* but not in *pe38* (Krappa et al., 1992; Kogan and Blissard, 1994). As observed in GATA, the mutation of the *cis*-acting CACGTG motif in UAR, a host factor binding sequence, eliminated host factor binding and reduced *gp64* promoter activity in the mutagenesis assay (Kogan and Blissard, 1994). This reflects a correlation between host factor interaction and transcriptional stimulation. Until now, four very early genes have been identified as transregulators: *ie-1, ie-0, ie-2* and *pe38* (Friesen, 1997). The product of *ie-1* (IE-1) is the principle transactivator of baculovirus gene expression. It stimulates transcription of promoters for the *39k, p35, gp64, p143, dnapol, pe38, lef-1, lef-2* and *lef-3* genes and also its own promoter (Pullen and Friesen, 1995; Ribeiro *et al.*, 1994; Lu and Carstens, 1993; Passarelli and Miller, 1993; Guarino and Smith, 1992; Blissard and Rohrmann, 1991; Kovacs *et al.*, 1991; Theilmann and Stewart, 1991; Nissen and Friesen, 1989). Moreover, IE-1 negatively regulates expression of *ie-0, ie-2* and *p38* genes (Kovacs et al., 1991; Leisy *et al.*, 1997). In addition, immediate early gene 1(*ie-1*) is indirectly and directly involved in DNA replication by stimulating late expression factors (*lefs*) and binding to *hr* elements which are thought to serve as origins of viral DNA replication (Kool *et al.*, 1993; Lu and Miller, 1995; Lu et al., 1997). IE-0 is thought to be the only baculovirus gene that can be spliced. The AcMNPV IE-0 transactivates early *39K* only in the presence of the *cis*-linked enhancer (Kovacs et al., 1991), while the OpMNPV and LdMNPV IE-0s can activate expression of the gene (*39K*) in the enhancer-dependent or -independent manner, although in the latter case the transactivation level from a *39K*-GUS reporter cassette, mediated by the IE-0, is comparatively low in the absence of the *hr* enhancer (Pearson and Rohrmann, 1997; Theilmann et al., 2001). The AcMNPV IE-0 stimulated expression

https://etd.uwc.ac.za/

from *ie-1* gene but was not autoregulatory (Kovacs et al., 1991), while the OpMNPV IE-0 can autoregulate its own promoter (Theilmann et al., 2001). IE-2 transactivates *39k, ie-0* and *ie-1* genes, although IE-1 negatively regulates *ie-2* (Leisy et al., 1997) and autoregulates its own promoter (Carson et al., 1991; Theilmann and Stewart, 1993). PE38 stimulates transcription of *p143* gene (Lu and Carstens 1993) and augments apoptosis induced by IE-1 (Prokhod'ko and Miller, 1999).

Unlike the early gene expression prior to the initiation of viral DNA replication, the late gene expression occurs during and after DNA replication. A virus-encoded RNA polymerase is responsible for transcription of late and very late genes. The late and very late promoter classes contain a consensus A/G/T(TAAG) sequence that serves as the late transcription initiation site. Only limited sequence beyond a few nucleotides surrounding the TAAG motif is required for late promoter activity (Morris and Miller, 1994; Ooi *et al.*, 1989; Rankin *et al.*, 1988). Eighteen late expression factors (*lefs*), *ie-1 ie-2, lefs 1-11, dnapol, p143, p43, p35* and *39K*, were previously identified (Lu and Miller, 1995). They are required for optimal transactivation of expression from late (*vp39*, major capsid protein, and *p6.9*, the basic core protein) and very late (*polh* and *p10*) promoters (Todd et al., 1995, 1996; Lu and Miller, 1994; Morris et al., 1994; Passarelli et al., 1994; Li et al., 1993; Passarelli and Miller, 1993a,b, 1994). Deletion of any of the *lef* clones from the *lef* library in transient expression assay dramatically reduced expression level from the *vp39* promoter (Lu and Miller, 1995). Recently, a 19th late expression factor (*lef-12*) has been reported to be necessary for expression of a reporter gene under control of late *vp39* and very late *polh* promoters (Rapp, et al., 1998). In addition, each of four *lefs*, *p47, lef-4, lef-*

https://etd.uwc.ac.za/

8 and *lef9* were identified to be one of four subunits of the virus-encoded DNA-direct RNA polymerase (Guarino et al., 1998). Apart from 19 *lefs*, very late expression factor-1 gene (*vlf-1*) was observed in transient assay to stimulate expression from very late *polh* and *p10* promoters, but has little or no effect on the late *vp39* and *p6.9* promoters (Todd et al., 1996). This suggested that *vlf-1* may be the major transactivator of very late gene expression. The two very late virus proteins, polyhedrin and *p10,* are expressed during the process of OB formation and ultimately ends at cell lysis. Their strong promoters are extensively utilized in baculovirus insect cell expression system (Smith *et al*., 1983; Martens *et al*., 1995).

The six *lefs, ie-1, lef 1-3, p143* and dnapol are essential for transient DNA replication (Kool et al., 1994) and the *lef 11* was also reported to be one of the essential genes for viral DNA replication (Lin and Blissard, 2002). Five other *lefs, ie-2, lef-7, p35, dnapol* and *pe38* stimulated transient DNA replication (Todd et al., 1995; Lu and Miller, 1995; Ahrens et al., 1996; Kool et al., 1995; Lu et al., 1997). Some replication-associated *lefs,* such as *danpol, lef-3* and *p143, are* directly involved in viral DNA replication, and some are probably indirectly involved in the replication process, such as *ie-2* and *pe38,* by stimulating expression of other *lefs* (Lu et al., 1997), while IE-1 is indirectly and directly involved in viral DNA replication by transactivating *lef-3, lef-7, p147, dnapol, p35* and *ie-2* promoters and binding to origins (*hrs*) of viral DNA replication (Lu and Miller, 1997).

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Homologous regions (*hrs*) are present in all the baculovirus genomes. Evidence suggestes that *hrs* act as enhancers to stimulate transcription of early genes in *trans* and *cis* (Guarino and summers, 1986; Guarino et al., 1986; Nissen and Friesen, 1989; Carson *et al*., 1988; Lu and Carstens, 1993), and serve as origins (*oris*) of DNA replication (Kool *et al*., 1993a,b; Pearson *et al*., 1992). They also indirectly contribute to late gene expression (Lu and Miller, 1995). IE-1 binds to the *hr* through interaction with the 28-bp imperfect palindrome. This is the minimal sequence required for IE-1-mediated *hr* enhancer activity (Rodems and Friesen, 1995) and plasmid DNA replication (Pearson et al., 1992; Leisy et al., 1995). Current evidence indicates that IE-1 interact with the 28-mer as dimer (Rodems and Friesen, 1995; Rodems et al., 1997). The IE-1 oligamerization and DNA binding are required for its transactivation function (Olson et al., 2001) and the binding of IE-1 to the *hr* may directly contribute to DNA replication (Lu et al., 1997).

1.6 Genetic Engineering and Expression Vectors of Baculoviruses

In nature the wild types of baculoviruses have some disadvantages of being commercialized. One of the major limitations is relatively slow speed of action, typically far less virulent in later star insects which cause the majority of damage to crops (Evans, 1981). A relatively narrow host range is also a drawback. Genetic engineering of baculoviruses may improve baculovirus properties as insecticidal agents for pest control.

Many studies have been conducted to enhance recombinant baculovirus insecticidal efficacy. Two main approaches are applied to explore the possibilities. Deletion of the viral *egt* gene (an ecdysteroid UDP-glucosyltransferase), the function of which is to either delay or block insect molting and pupation, and the introduction of a scorpion toxin

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

gene and a mite toxin gene can increase viral killing speed to some extent. (O'Reilly and Miller, 1991; Slavicek *et al.*, 1999; Stewart *et al.*, 1991; Maeda *et al.*, 1991; McCutchen *et al.*, 1991; Gershburg *et al.*, 1998; Tomarski and Miller, 1991; 1992). However, over-expression of insect genes in recombinant baculoviruses, such as a modified juvenile hormone esterase gene (JHE), did neither significantly increase the baculovirus speed of action nor reduce the food consumption of the infected insects (Hammock *et al.*, 1990). The different promoters, either from a powerful late polyhedrin and a p10 promoter or an early promoter, have been studied for improvement of the insecticidal activity (King and Possee, 1990; O'Reilly *et al.*, 1992; Luckow, 1991; Jarvis *et al* 1996a; 1996b). The effectiveness of the genetically modified baculoviruses has also been tested in the field and their potential to enhance crop protection has been reported (Cory *et al.*, 1994; Gard, 1997; Hernaandez-Crespo *et al.*, 1999). The AcMNPV and HaSNPV recombinants with an *egt*-deletion and an insertion of the insect-specific neurotoxin (*Aa*IT) had shown a faster killing speed over their wild types and the *egt*-deletion forms respectively. (Black *et al.*, 1997; Chen *et al.*, 2000;). The recombinants thus exhibit a possible commercial potential as bio-insecticides.

UNIVERSITY of the
WESTERN CAPE

Other application of baculoviruses is to produce recombinant proteins for biotechnological and pharmaceutical interests by baculovirus expression vectors (Possee, 1997; Krost and Condreay 1999). The baculovirus-derived vectors have also been used to transfer genes into mammalian cells (Hofmann *et al.*, 1995; Krost and Condreay, 1999; Sarkis *et al.*, 2000) and it thus could be a possible tool for gene therapy. The polyhedrin or p10 locus are usually adopted to introduce the foreign genes as both genes have a very

https://etd.uwc.ac.za/

strong promoter and are not essential for infectivity of NPVs in cultured insect cells (King and Possee, 1992; O'Reilly *et al.*, 1992; Luckow, 1991).

1.7 *Trichoplusia ni* and TnSNPV

Trichoplusia ni (*T.ni*) (Lepidoptera: Noctuidae) is commonly referred to as cabbage looper. The average generation time of *T.ni* is between 26-40 days. Following mating, females lay between 200-300 eggs on the upper or lower surfaces of upper canopy leaves. Once eggs hatch, the pale green larvae feed actively on veins on the underside of lower leaves for 14-21 days, after which they pupate. Nine days later, after initial pupae develop, an adult moth emerges. Mature moths feed on plant nectar and can live for up to 24 days (Adlerz, 1971; Kishaba *et al.*, 1976; Marsden, 1979; Hofmann and Fordsham, 1993).

T.ni is a polyphagous crop pest distributed throughout Africa, North America, Hawaii, Europe and Asia. The larvae cause damage on cabbage, spinach, sugarbeets, peas, celery, potatoes, alfalfa, beans, tomato, certain ornamental plants and mint, etc. Although *Trichoplusia ni* has been successfully controlled with synthetic chemical pesticides, increased resistance to these insecticides such as *Bacillus thuringiensis* (Bt) has been reported (Dornan *et al.*, 1995). The combined use of traditional pheromone releasers and black light traps has proven inadequate in crop protection (Debolt *et al.*, 1979). Thus, an effective long-term pest control agent is urgently needed.

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

Single nucleocapsid polyhedrosis virus is one of morphotypes of baculoviruses. In this morphotype viruses are packaged as a single nucleocapsid within each viral envelope. To date, several single nucleocapsid polyhedrosis viruses were either fully sequenced or under investigation. Since TnSNPV was cultured *in vitro* by Granados *et al.* (1986), a TnSNPV isolate has been characterized by Davis and Wood (1996). In recent years a South African TnSNPV isolate was isolated from a field population of *Trichoplusia ni* and partially characterized (Fielding, Ph.D. thesis, 2001). Although a previous attempt to construct a physical map of the entire TnSNPV genome was unsuccessful, an approximate genome size of 158 kb was determined by restriction endonuclease analyses (Fielding, Ph.D. thesis, 2001). The size is about 40 kb greater than those of the previous TnSNPV isolates (Bilimoria, 1983; David and Wood, 1996). Five genes within the TnSNPV had been fully sequenced. A few attempts made to phylogenetically define the relation of the virus to other baculoviruses have suggested that the TnSNPV belong to group II NPVs (Wang *et al.*, 2001; Fielding and Davison, 1999; Fielding *et al.*, 2002).

1.8 Scope of the Thesis

UNIVERSITY of the
WESTERN CAPE

Along with the development of biological insecticides, many research efforts have been made in baculoviruses to investigate fundamental molecular aspects of these viruses, such as the function and regulation of genes, genome organization, mode of entry, DNA replication and virus factors that determine the host range and virulence. Previously, a South African TnSNPV isolate was partially characterized as a novel baculovirus. During the process of the characterization, a few late genes of the virus were identified. This thesis describes a molecular characterization of the TnSNPV early genes to gain insight

https://etd.uwc.ac.za/

into the functional roles of these genes, their unique features and further determination of the placement of TnSNPV in baculovirus phylogeny.

In the cascade of baculovirus regulatory events, successive stages of virus replication are dependent on proper expression of genes within the preceding stages. Thus, critical to baculovirus replicative success is the appropriate expression and regulation of early genes (Friesen, 1997). Immediate early gene 1 (*ie-1*) is thought to be a principal transregulator, and conserved among baculoviruses. It potently stimulates transcription of promoters for some early and late genes as well as its own promoter. *Ie-1* gene is also directly or indirectly involved in DNA replication. The N-terminal half of IE-1 has been identified as a transactivating domain which has an acidic feature and is poorly conserved. The C-terminal half of the gene has a DNA-binding function and is comparatively better conserved. The sequence analyses of the TnSNPV *ie-1* gene will provide useful information for further understanding of the essential regulatory gene, as well as for phylogenetic classification of the TnSNPV due to its advantages over polyhedrin and other conserved late genes with small sizes which were previously used as models for phylogenetic study (**Chapter 2, 4**).

Current evidence suggests that IE-1 transactivation involves DNA-independent and *hr* enhancer-dependent mechanisms (Friesen, 1997). In a DNA-independent model, IE-1 interacts directly or indirectly with a transcriptional initiation complex without involvement of DNA binding in a sequence-specific manner. In a *hr* enhancer-dependent model the binding of IE-1 to the *hr* occurs through interaction with the 28-mer

https://etd.uwc.ac.za/

palindromes within the *hr,* and IE-1 binds to the *hr* enhancer element as a dimer which is a requirement for IE-1 transregulatory function. The binding stimulates *hr* enhancer activity and may lead to transcriptional stimulation through direct and indirect protein contact with TFIID-associated factors of the preinitiation complex (Friesen, 1997). Although deletion and insertion studies on IE-1 binding have suggested that deletion of more than 25 amino acids at the C-terminus demolished the IE-1 DNA-binding capacity (Kovacs et al, 1992), and a few C-terminal residues are responsible for oligomerization of IE-1 (Rodems et al., 1997; Olson et al., 2001), which regions or domains of IE-1 directly bind to the *hr* enhancer element remains to be determined. Based on conservation of *hrs* in all the lepidopteran baculovirus genomes, it was thought that TnSNPV may have *hr* sequence homologs as presented in other baculovriuses. However, due to lack of entire TnSNPV genome information, the TnSNPV *hr* sequences or other possible substitutes of enhancer elements of the TnSNPV IE-1 could not be determined, and therefore it was decided to use AcMNPV IE-1 as a model to experimentally predict the existence of the possible direct binding region in the TnSNPV IE-1, based on conserved primary and secondary structures between both protein analogs (**Chapter 3**). This study will contribute to the further understanding of the functional role of IE-1 in DNA binding.

Another early gene was previously identified as a major early-transcribed gene. Its homolog was found in all the available baculovirus genomes. Although it was suggested that ME-53 may play a role in gene regulation, based on the presence of a zinc finger-like motif in its amino acid sequence, the regulatory role of the gene is unclear. Thus, identification and characterization of the TnSNPV *me-53* early gene will enrich

https://etd.uwc.ac.za/

knowledge we have obtained only from its AcMNPV counterpart and give a better understanding of the common features around all the *me-53*s as well as the unique features of the TnSNPV *me-53*. In addition, the comparative phylogenetic analysis of ME-53s and IE-1s provides information for further defining phylogenetic relationships of TnSNPV among baculoviruses.

The second goal of this study is to test multiple foreign gene expressions under the control of heterologus early promoters in the context of AcMNPV genome. Although an insect-specific toxin gene (AaIT) under the control of a baculovirus early promoter was tried to improve the properties of baculovirus recombinant as a bio-control agent (Jarvis et al., 1996), multiple early expressions of insect-specific toxin genes have not been reported. The introduction of multiple toxins into baculoviruses, expressed early in infection may reduce insect feeding and, finally, result in feeding arrest or death quicker than those induced in the very late phase of infection, because different toxins may act on the different physiological pathway of larvae development and/or the same multiple toxins may be rapidly accumulated to a killing threshold dosage. However, baculovirus native early promoters are comparatively weaker than very late ones and therefore overexpression level of a toxin gene may be much lower than those which are under control of late strong promoters, such as *polh* and *p10* promoters. Selection of comparatively strong early promoters, such as *ie-1* or *gp64* promoters, can be taken into account. Associated with the promoter selection, enhancer elements can be applied to facilitate overexpression of selected early promoters. With respect to enhancement of early promoters, several enhancer-dependent early genes can be chosen for this purpose.

https://etd.uwc.ac.za/

In addition, appropriate heterologous early promoters, such as enhanced CMV promoter (Clontech) or even hybrid early and late promoters, to drive expression of insect-specific toxins, can also be candidates in this regard. However, these possibilities still need to be tested in both laboratory and field.

It was reported that very late gene expression does not occur until all the virus-encoded factors become available during a very late phase of infection (Li, et al., 1993; Passarelli and Miller., 1993a,b,c; McLachlin and Miller, 1994; Passarelli, et al., 1994; Morris et al., 1994; Todd et al., 1995; Xu et al., 1995), and it is also undesirable for very late promoters to produce secreted proteins and glycoproteins because it appeared that cellular processing pathways may be compromised during the late stage of infection (Jarvis and Summers, 1989; Jarvis et al., 1990; Murphy et al., 1990). Although the expression of toxin genes from very late promoters appears to have such disadvantages for pest control purposes, the overexpression of insect-specific toxin genes controlled by strong very late promoters, *polh* or *p10*, showed a dramatic increase of killing speed of infected insect hosts which was faster than that of the wild type virus (McCutchen et al., 1991; Chen, et al., 2000). However, the recombinant virus carrying a single toxin gene under control of an immediate early promoter did not show promising results, although it showed the reduction of insect growth rate (Jarvis et al., 1996),. The main reasons for this may be summarized as 1) a threshold dose for killing was not reached due to the accumulation of the toxin expressed by the early promoter which was slower and less than that of the very late promoters, specifically at the very late postinfection; 2) the duplication of the early promoter may slow down the viral replication cycle because the duplicate of the promoter

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

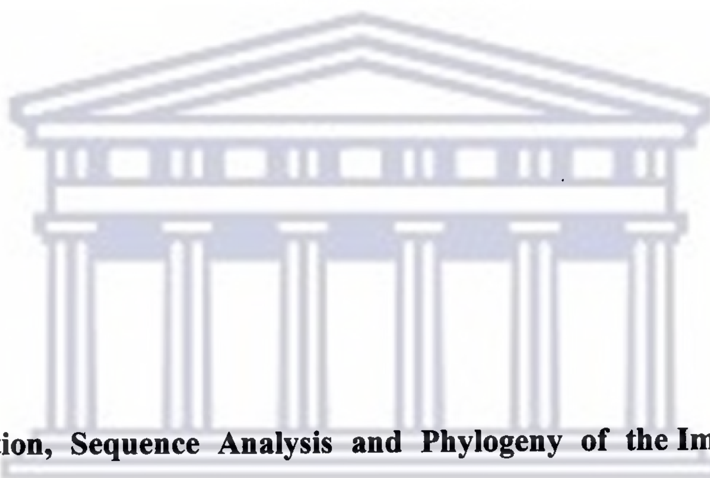
competes limited cellular transcription factors with its native counterpart; 3) the expression of the toxin (AaIT) in the midgut epithelial cells could impede efficient virus spread to other tissues in the infected insect host (Jarvis et al., 1996). Although expression of a single toxin gene under the control of an early promoter could not produce as much active impact on enhancing baculovirus recombinant killing speed of infected insects, as demonstrated in the case of a very late promoter, the strategy for early expression of multiple toxin genes may bring encouraging views into the challenging topic.

Due to insertion of two heterologous genes into the baculovirus expression vector system, the expression of these genes in the recombinant virus will provide evidence for developing a model system, possibly used for introduction of multiple insect-specific toxin genes under the control of early promoters without interfering with virus replication (**Chapter 4**). Thus, virulence and killing speed of the virus as a bio-insecticide may be dramatically boosted and improved.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Chapter 2



Identification, Sequence Analysis and Phylogeny of the Immediate Early Gene 1 of the *Trichoplusia ni* Single Nucleocapsid Polyhedrosis Virus

https://etd.uwc.ac.za/

2.1 Abstract

The immediate early gene 1 (*ie-1*) as a principle transactivator protein (IE-1) plays an essential role in gene regulation and DNA replication in the case of of *Autographa californica* multicapsid nuclear polyhedrosis virus (AcMNPV) and the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus (OpMNPV). In this study an *ie-1* homolog from *Trichoplusia ni* single nucleocapsid polyhedrosis virus (TnSNPV) was identified. It has a comparatively larger open reading frame (ORF), encoding a protein of 739 amino acids with a molecular mass of 84.5 kDa. Each promoter region of compared NPV ie-1s contains the TATA and CAGT/T motifs. Protein analysis showed that the N-termini of IE-1s are poorly conserved, but rich in acidic amino acids. Two negative charge dominant regions were identified in each N-terminus. The SwissProt database search showed that N-terminal region of the TnSNPV IE-1 significantly matched the SET-domain. On the other hand, the C-terminal regions of the deduced IE-1 proteins were much more conserved than the N-termini. Several conserved regions were found in C-termini. The TnSNPV *ie-1* shared the highest sequence homology with the *Spodoptera exigua* MNPV (SeMNPV) homologue in comparison with others. The phylogenetic analysis showed that TnSNPV is a member of the group II NPVs and it is most closely related to SeMNPV.

https://etd.uwc.ac.za/

2.2 Introduction

Baculoviruses are a diverse group of insect viruses with circular double stranded DNA genomes ranging in size from 90kb to 180 kb (Herniou *et al.*, 2003). The most well-characterized baculovirus, AcMNPV encodes approximately 150 genes (Ayres, 1994). Gene expression of baculoviruses are transcriptionally regulated in a cascade manner, which includes three temporal phases: early, late and very late (Friesen, 1997). Early gene expression precedes viral DNA replication, dependent on host RNA polymerase II and does not require the presence of any viral gene products, while late and very late genes are transcribed by a virus-encoded RNA polymerase during or after the initiation of viral DNA synthesis (Blissard and Rohrmann, 1990; Guarino *et al.*, 1998). Some early genes such as *ie-0*, *ie-1*, ie-2 and *pe38* play critical roles in the regulation of viral gene expression (Frisen, 1997).

The *ie-1* gene encodes a principal early transregulator protein (IE-1) and expresses at a very early stage of infection and remains detectable throughout infection (Frisen, 1997; Theilmann and Stewart, 1990). The early transcription of the AcMNPV, OpMNPV and LdMNPV *Ie-1*s was initiated from a baculovirus consensus early promoter motif CAGT (Theilmann and Stewart, 1990; Pullen and Friesen, 1995; Van Strien *et al.*, 2001), but in the case of the AcMNPV *ie-1* the motif can also function as a late promoter (Pullen and Friesen, 1995). IE-1 stimulates the transcription of some early, late and its own promoters (Pullen and Friesen, 1995a; Ribeiro *et al.*, 1994; Lu and Carstens, 1993; Passarelli and Miller, 1993; Guarino and Smith, 1992; Blissard and Rohrmann, 1991; Kovacs *et al.*,

https://etd.uwc.ac.za/

1991; Nissen and Friesen, 1989), and also negatively regulates *ie-0, ie-2* and *pe38* promoters (Kovacs *et al.*, 1991; Leisy *et al.*, 1997). IE-1 mediates the transcriptional activation of early promoters by enhancer-dependent and -independent mechanisms (Guarino and Summers, 1986b,c; 1987; Guarino *et al.*, 1986a; Nissen and Friesen, 1989; Blissard and Ro*hr*mann, 1991; Blissard *et al.*, 1992). In the former when homologous region (*hr*) enhancer elements were cis-linked to the upstream of early promoters (*39K*, *p35* and *p143*), expression of reporter genes was dramatically stimulated by IE-1, while in the latter case the early promoters (*39K* and *gp64*) were transactivated by IE-1 in the absence of *hr* enhancer elements. *Ie-1* is one of five genes that are essential for transient DNA replication (McDougal and Guarino, 2000). It is indirectly and directly involved in viral DNA replication by transactivating some *lefs* and binding to *hr* regions (Lu and Miller, 1997; Lu *et al.*, 1997). In addition, IE-1 alone is sufficient to induce apoptosis in Sf21 cells, the function of which can be augmented by the involvement of *pe38* (Prikhod'ko and Miller, 1996; 1999).

UNIVERSITY of the
WESTERN CAPE

It is known that *ie-1* is the only baculovirus gene that can be spliced. Its spliced form is IE-0 which differs from IE-1 by addition of different numbers of amino acids to the N-termini of different baculovirus IE-1s (Kovacs *et al.*, 1991; Pearson and Ro*hr*mann, 1997). Unlike AcMNPV and OpMNPV IE-1s essential for both transactivation and DNA replication, the LdMNPV IE-0 showed the dual functions in the Ld652Y cell line but the LdMNPV IE-1 was inactive. This suggested that splicing in the case of the LdMNPV IE-1 is important for producing an active transactivator that has the same functions as those of other unspliced IE-1s (Pearson and Ro*hr*mann, 1997). It is possible that IE-1 and IE-0

https://etd.uwc.ac.za/

are translated from the LdMNPV *ie-0* transcripts as in the case of the OpMNPV *ie-0* (Theilmann *et al.,* 2001). The function of the LdMNPV *Ie-1* has not been identified.

The previous studies on IE-1s showed that the amino acids at the N-terminal half of IE-1s are poorly conserved and rich in acidic features, while the C-terminal residues of IE-1s are more conserved (Theilmann and Stewart, 1990; van Strien *et al*., 2000). Several domains have been identified as being involved in transactivation and DNA replication at the N-termini of IE-1s. One or two acidic transactivation domains were mapped in the different regions of N-termini of the AcMNPV or OpMNPV IE-1s, respectively (Slack and Blissard, 1997; Rodems *et al*., 1997; Forsythe *et al*., 1998), whereas in the case of the AcMNPV IE-1s it was suggested that two N-terminal regions with basic charges may be possibly involved in negatively regulating transcription from *ie-0, ie-2* and *pe38* (Slack and Blissard, 1997). More recently, amino acids 1 to 65 within the N-terminal acidic activation domain of OpMNPV IE-1 were identified to be essential for DNA replication but functionally separated from the transcriptional activation domain (Pathakamuri and Theilmann, 2002). The C-terminal half of IE-1 contains a few residues and a helix-loop-helix-like domain which are critical for IE-1 oligomerization and DNA binding. The binding of IE-1 to *hr* enhancer elements as a dimer is required for transactivation of early promoters (Rodems *et al*., 1997; Olson *et al*., 2001).

Previously, based on the morphology of the occlusion bodies, baculoviruses were classified into two genera: nucleopolyhedrosis virus (NPV) and granulosis virus (GV) (Murphy *et al*., 1995; Blissard *et al*., 2000). The phylogenetic relatedness of lepidopteran

https://etd.uwc.ac.za/

NPVs have been further divided into group I and II based on a single gene (*polh*) phylogeny (Zanotto *et al.*, 1993). The analyses from a number of individual gene phylogenies also contributed to the classification (Chen *et al.*, 1999; Clarke *et al.*, 1996; Jin *et al.*, 1999; Kang *et al.*, 1998; Li *et al.*, 2001; Liu *et al.*, 1999). In the recent years with increase of available genome sequences, the classification has been clarified and analyzed in details by several combined analyses based on baculovirus genome information. Most recently, a dipteran NPV (*Culex nigripalus* NPV, CuniNPV) has been identified and fully sequenced (Moser *et al.*, 2001; Afonso *et al.*, 2001). Herniou *et al.* (2003) suggested that baculoviruse phylogeny can be reconstructed to four major groups based on a number of major differences of CuniNPV from other lepidopteran baculoviruses (Afonso *et al.*, 2001). In spite of the major change of baculovirus classification, the phylogenies of group II NPVs are not well resolved because of a lack of sufficient sequence data available in the divergent group. The sequence data of South Africa TnSNPV isolate will provide useful information to contribute to baculovirus phylogeny. Previous studies suggested that the TnSNPV is more closely related to group II NPVs based on phylogenies of a few late genes (Fielding and Davison, 1999; Fielding *et al.*, 2002). However, phylogenetic relationship in deep branches among TnSNPV and other baculoviruses has not yet been established. Identification of the TnSNPV *ie-1* homolog, which is conserved among all the lepidopteran baculoviruses, will contribute to the TnSNPV phylogeny.

In this study the TnSNPV *ie-1* nucleotide and deduced protein sequences were analyzed and discussed in detail. The eleven baculovirus IE-1 sequences were compared to identify

https://etd.uwc.ac.za/

common features of the early regulatory gene. The phylogenetic analysis of TnSNPV has also been conducted for further establishment of the virus phylogeny among baculoviruses.

2.3 Materials and Methods

2.3.1 Insects, virus isolation and purification

The original virus isolate was prepared from diseased *Trichplusia ni* (Noctuidae: Lepidopteran) larvae collected from the Eastern Cape, South Africa (Fielding and Davison, 1999). Third star Larvae were fed on small plugs of an artificial lepidopteran diet, containing 1 μl purified polyhedral inclusion bodies (PIBs) per piece. After the entire diet plug was consumed, each of the infected larvae were transferred to fresh diet in glass containers and reared at 26°C and 65% humidity on a 12 hour day/night cycle until death. Polyhedral inclusion bodies (PIBs) were purified from virus-killed larvae following a modified procedure based on the previously described method (Miller and Dawes, 1978). Briefly, ten collected insect cadavers were homogenized in 10 ml H_20 within a pestle using a mortar, and the mixture was filtered through two layers of muslin cloth into a beaker. Debris was removed by centrifugation at 2000 RPM for 5 min. The supernatant was centrifuged at 15000 RPM for 10 min. The pellet was resuspended in 10 ml 0.1% SDS and centrifuged at the same speed for 10 min, and the step repeated with the same volume of 0.5M NaCl. The following pellet was resuspended with gentle shaking in 5 ml H_20 overnight. The 1.2 ml preparation was loaded onto a discontinuous sucrose gradient (40, 45, 50, 55 and 60 %) and centrifuged at 26500 RPM for 3 hrs.

https://etd.uwc.ac.za/

Visible band at 55% sucrose layer was collected and diluted up to 25 ml with H_20. PIBs were centrifuged at 15000 RPM for 10 min. The pellet was resuspended in 1 ml H_20 and store at 4°C.

2.3.2 DNA extraction, manipulation and sequencing

TnSNPV DNA was isolated from the occluded form of the virus (Fielding and Davison, 1999). PIBs (2 x 10^6 PIB/ml) were placed in 0.5M Na_2CO_3 alkaline solution at 37°C for 30 min. SDS was added to the solution to a final concentration of 0.1%, and PIBs were incubated at 60°C for 30 min. After cooling down, PIBs was digested by proteinase K with a final concentration of 500 μl/ml at 37°C for 1 hour (or until clear), followed by phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform-isoamyl alcohol (24:1) extractions. Viral DNA was recovered by ethanol precipitation and the concentration of DNA was determined by measuring the OD_{260} and OD_{280} of the solution.

The virus genomic DNA was used to construct genomic libraries with *Pst*I and *Eco*RI. Ten of the 29 *Pst*I and 22 of 27 *Eco*RI fragments were cloned into pSK-Bluescript (Stratagene) (Fielding, 2001). As part of a preliminary attempt to map the genome, the ends of the *Eco*R1 library were sequenced. Subsequent genome analysis led to the identification of the TnSNPV homolog. The *Ie-1* gene was found to be truncated with its 5' and 3' ends on a 2.3 kb and an 11 kb *Eco*RI fragment, respectively. The 11 kb and 2.3 kb DNA fragments were digested by exonuclease III (Henikoff, 1984) and then cloned into pSK-Bluescript vector. Appropriate sizes of fragments were selected and DNA templates were prepared for nucleotide sequencing. Sequencing was conducted using the

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Sequitherm kit (Epicentre Technologies) using CY-5 labeled primers. Nucleotide sequence was resolved on an Alfexpress automated DNA sequencer (Pharmacia). Sequence was obtained in both the sense and anti-sense directions before the final sequence was confirmed (Fig. 2.1).

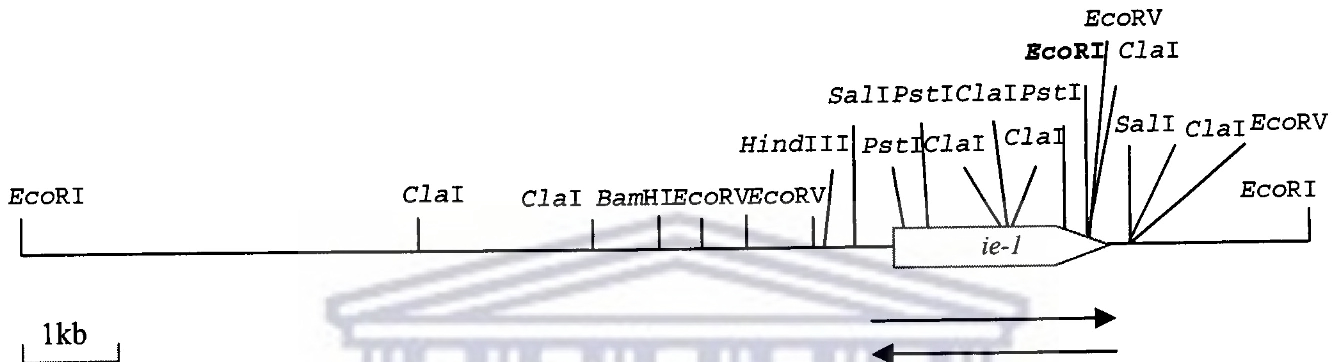


Figure 2.1 The location of the *ie-1* gene within an 11 kb *Eco*RI fragment and an adjoining 2.3 kb *Eco*RI fragment. An *Eco*RI restriction site between the 11 kb and 2.3 kb fragments is in bold. The orientation of the gene and sequencing directions are indicated by arrows.

UNIVERSITY of the WESTERN CAPE

2.3.3 Computer analysis

Nucleotide and amino acid sequence manipulation was carried out using the University of Wisconsin, Genetics Computer Group (GCG) sequence analysis package. The BLAST algorithm of Altschul *et al.* (1990) was used to compare sequences generated in this study with entries in non-redundant nucleotide and protein sequences databases accessed by the National Center for Biotechnology Information (NCBI). IE-1s were analyzed using the ExPASY server (Appel *et al.*, 1994) for charge analysis and domain prediction. Multiple

https://etd.uwc.ac.za/

sequence alignments were conducted using the ClustalW program of Thompson *et al.* (1994). The alignment was used as the input to construct the phylogenetic trees by using neighbor-joining method implemented in the Clustal W program and the robustness of the data sets was estimated using bootstrap resampling procedure. GenDoc software was used for similarity shading and scoring among the aligned sequences.

2.4 Results

2.4.1 Nucleotide and protein sequence analyses

The TnSNPV *ie-1* gene has an open reading frame (ORF) of 2217 nucleotides, encoding a protein of 739 amino acids, with a molecular mass of 84.5 kDa (Fig. 2.2). The ORF is slightly larger than other compared *ie-1*s in this study which range from 1578 to 2142 nucleotides for McNPV and SeMNPV species respectively.

```
1      TTAACTCGGTCAGCGCGCCCGCGTTCGTAGCCGCTTACGAGTCTTGTGGAAGCGAAATTG

61     CGGTTATTACCATCGCCGTCAGTGTCATCGCTTTCGTTTATTTGAAACGGTATGATATTT

121    ATCGAGTTGATCGCATTTATCACGCTAAATGTAATTAGGCTATCTTTATCTTGAGGATAG

       -55                           -24
181    TATAAATTGACAATCATTTTATTTTTACTTCCAGTTCTCACTTACAAGCGTTGTGATGTC
                                                                M  S   2

241    GCATCCAAACATCATCAGTGCCATGGACAACAACGACGCGTCTATCAAATATAAAAATTA
         H  P  N  I  I  S  A  M  D  N  N  D  A  S  I  K  Y  K  N  Y   22

301    TATCGACAACGCCATCAACACTCCTACGCATACGATTCTGCAGAACGTCAGTATGGATTT
         I  D  N  A  I  N  T  P  T  H  T  I  L  Q  N  V  S  M  D  F   44

361    CGACGACAGCAATATTCTGGATTTCGGCAACGAAAATGATATGAATGTCTACGACAGACG
         D  D  S  N  I  L  D  F  G  N  E  N  D  M  N  V  Y  D  R  R   64

421    AGACAGTAACAGCAGTAAAATTGTGAATGATGCTTGCGATGAAAACTCTCAATTTTCTGA
         D  S  N  S  S  K  I  V  N  D  A  C  D  E  N  S  Q  F  S  D   84
```


https://etd.uwc.ac.za/

```
481   TGTCAACGTCAACAATAATGCGGACAACAATGATTATATAAAAATCATGAAAACTGCTAC
       V  N  V  N  N  N  A  D  N  N  D  Y  I  K  I  M  K  T  A  T   104

541   CGATGTCGTCGAAAACAAGAATGAATATACGAATAAACATAAAACTGCAGTCGTTTCGAC
       D  V  V  E  N  K  N  E  Y  T  N  K  H  K  T  A  V  V  S  T   124

601   TAAACCATTCAAGAAAAATCCTAAAAAAAGGCCATCGTCATCGTTGACGACTACGACGAC
       K  P  F  K  K  N  P  K  K  R  P  S  S  S  L  T  T  T  T  T   144

661   GACGACGACAGAAAAGAAGAACAAGTCAAGACCCAACAGACCTCCCAATTCCACTGTAAT
       T  T  T  E  K  K  N  K  S  R  P  N  R  P  P  N  S  T  V  I   164

721   CGCTGATGGTAGTATTCCACCACAACCTGTGATAAAGCCATCGAAGAAGCAGACAGTTTT
       A  D  G  S  I  P  P  Q  P  V  I  K  P  S  K  K  Q  T  V  F   184

781   TGTTTCGCCTTTAATCAATAGAGGAGGAAAAAACTTGAATGTGTTGCGCAACGACAATAA
       V  S  P  L  I  N  R  G  G  K  N  L  N  V  L  R  N  D  N  N   204

841   TAATAATTTCAATAATGACAGTGACGATAGCAACGGAAGCGACAGTGAAGATAGCGATTC
       N  N  F  N  N  D  S  D  D  S  N  G  S  D  S  E  D  S  D  S   224

901   TACGCATCCGCCGCCTTCGAAAAAGACAAAAATGACATCAAAATCATCTAAAATGTCGGT
       T  H  P  P  P  S  K  K  T  K  M  T  S  K  S  S  K  M  S  V   244

961   GACGCCGCAACAACAAATGCCCGAGATTTTGAAAATTAATGCTGCCGACAAAAATAAAGT
       T  P  Q  Q  Q  M  P  E  I  L  K  I  N  A  A  D  K  N  K  V   264

1021  CAATGACGAGAAACAAACAGTGAAATATAACAAAAAAAAGCAACAATCTCAAGACGCTGG
       N  D  E  K  Q  T  V  K  Y  N  K  K  K  Q  Q  S  Q  D  A  G   284

1081  TGCCGTCGTGGTCGTGAAACAACAAAAACTTGATAACGAATCAACAAGTCAAACTTCCGT
       A  V  V  V  V  K  Q  Q  K  L  D  N  E  S  T  S  Q  T  S  V   304

1141  TAATGATGATCAACAACGGTCGAAAGATTGCGATTCTCCAACAAATGACTTGTTTGAAAA
       N  D  D  Q  Q  R  S  K  D  C  D  S  P  T  N  D  L  F  E  N   324

1201  TAAAATAATCCCCAACATGATGACCATGGAAAGAGACAATAACCGCAAGTTTGTGCAATA
       K  I  I  P  N  M  M  T  M  E  R  D  N  N  R  K  F  V  Q  Y   344

1261  TATTCTCAACGCTCACAATTATCTGTTTATAGTATACGAAAACAAGTATAATGCCAAGAC
       I  L  N  A  H  N  Y  L  F  I  V  Y  E  N  K  Y  N  A  K  T   364

1321  TTTTAACAAAAACTCCAACGCATCGATTTATAAAATAGAGTATGTGAATTGCGTCCAGTC
       F  N  K  N  S  N  A  S  I  Y  K  I  E  Y  V  N  C  V  Q  S   384

1381  CATATACAAGTATTATAACGCCAATTACTCGCATATCGATAGAACATGCAAAGTCGTGTC
       I  Y  K  Y  Y  N  A  N  Y  S  H  I  D  R  T  C  K  V  V  S   404

1441  TTTCAATCGATTCAGATTCGCCATATCTGTGAACCTTTTAAATAAAATGCAGATTGAATT
      F  N  R  F  R  F  A  I  S  V  N  L  L  N  K  M  Q  I  E  L T  424

1501  GCCTCCTACGGAACAATTTAAAAAGGAAGACCTCAAGAAGATTTCTCCGAAGAACACTTT
       P  P  T  E  Q  F  K  K  E  D  L  K  K  I  S  P  K  N  T  F   444

1561  TTGCCTATTAAATGAAGTCAAAGATCCGGATTTCATTTCCAAGCTCACTAACACATTCGG
       C  L  L  N  E  V  K  D  P  D  F  I  S  K  L  T  N  T  F  G   464

1621  CTTGGACAATATTTATATTCAGGGTCAACTCACTATGCTGCTCTCGTCGATTGGTGAGAA
       L  D  N  I  Y  I  Q  G  Q  L  T  M  L  L  S  S  I  G  E  N   484

1681  TCGGGCAAAGATTTTGAATCAGCATATCAGTGCAATGATTGAAGATAAAAGCCTATTCAC
```


https://etd.uwc.ac.za/

```
          R  A  K  I  L  N  Q  H  I  S  A  M  I  E  D  K  S  L  F  T  504

1741  TATACCTTTGCATTTGTCTCGATCCAAGGAATTGGAAGAAATTGTCGATGACGATCTGAA
          I  P  L  H  L  S  R  S  K  E  L  E  E  I  V  D  D  D  L  N  524

1801  CCCCAACAACAGTAACGTCTCGTCGGCCTACATTCGAGACATAATAGAACTCTCGAACAA
          P  N  N  S  N  V  S  S  A  Y  I  R  D  I  I  E  L  S  N  K  544

1861  ACTCAAGTTTAAGGCTCCTATTATTCCGTCATATGTCCATAAAACCAAGGAACAAAACAT
          L  K  F  K  A  P  I  I  P  S  Y  V  H  K  T  K  E  Q  N  I  564

1921  TGAGAATGTTCTTAGTTTTTGGATCAACACTCAGAAGAACAACAACGAGCGCGATAAAAC
          E  N  V  L  S  F  W  I  N  T  Q  K  N  N  N  E  R  D  K  T  584

1981  TTTGGCAAAATCTCTGCAGTTTACATACAAGTTTACCAGTGTCGCTCGAGTGCTCTTCGA
          L  A  K  S  L  Q  F  T  Y  K  F  T  S  V  A  R  V  L  F  D  604

2041  CGAAACCGATGGCGACGTCAATAAACTGTTTAAAGTGAAGAAAGAGCCTGGATCTGTGGC
          E  T  D  G  D  V  N  K  L  F  K  V  K  K  E  P  G  S  V  A  624

2101  AATGATTGAAGATTATCTACAGGCTTGTGAAAAAATACCCAATGGCAACAACTTTATTAT
          M  I  E  D  Y  L  Q  A  C  E  K  I  P  N  G  N  N  F  I  M  644

2161  GATCAACACACTCAACGATGAACGCGTGACAATCATCAAGGCCAAAAATGAATTCTTTTG
          I  N  T  L  N  D  E  R  V  T  I  I  K  A  K  N  E  F  F  W  664

2221  GATTCGTACTAATAATCCTAATAATTTAATTCACTGTATCGATATCATTATGGCCTTTAA
          I  R  T  N  N  P  N  N  L  I  H  C  I  D  I  I  M  A  F  K  684

2281  AAACTTTAATCATCATTTGCTCTCTTTGATTCCCAGCAATCGTAAGGATTTGAACAATCG
          N  F  N  H  H  L  L  S  L  I  P  S  N  R  K  D  L  N  N  R  704

2341  TCACAGTGGATTAATAAAGCTAGTGGCCTATCATTTAGGTGGTGATGTTGACATTAATTT
          H  S  G  L  I  K  L  V  A  Y  H  L  G  G  D  V  D  I  N  F  724

2401  TGTACGTGCCATGGCTGAGAAGTTTAAATGTAATTATCTATATAAAAAATTTTAAATGTA
          V  R  A  M  A  E  K  F  K  C  N  Y  L  Y  K  K  F  *        739

2461  TAATTTTTCTTGTCTACTGTAATTTGAATAAATTTTTAAGGATCGTATTTTGTATAG
```

Figure 2.2 Nucleotide and deduced amino acid sequences of the TnSNPV *ie-1* gene. A putative TATA box and a CAGT motif are underlined. The positions of these elements in the *ie-1* promoter region are indicated by numbers. An *Eco*RI restriction site located at the truncated ends of two fragments (11 kb and 2.3 kb) and a single polyadenylation signal in the 3' non-coding region are double underlined. The SET-domain-like region of 47-120 amino acids at the N-terminus of the deduced IE-1 protein are in bold.

https://etd.uwc.ac.za/

A consensus baculovirus early promoter motif CAGT and a TATA box were found at -24 nt and -55 nt, relative to the predicted translational start codon of the TnSNPV *ie-1*, respectively. (Fig. 2.2). A comparison of the promoter regions of the eleven *ie-1* genes indicated that a TATA box is completely conserved among these regions, and a CAGT element is highly conserved with the exception of the HzSNPV and HaSNPV *ie-1* promoters, both of which have a CATT motif presented at the same position of -47, relative to its translation start codon (Fig. 2.3). Two CAGT motifs were found downstream of a conserved TATA box in the case of the CfMNPV and SpltNPV *ie-1s*, respectively. A single polyadenylation signal incorporating the TAA termination codon was found at the 3' non-coding region of the TnSNPV *ie-1* (Fig. 2.2).

```
AcMNPV  : -GGCATAATAAAGAAATATTGTTATCGTGTTCGCCATTAGGGCAG :  45
BmNPV   : GGGCATAATAAAAAAATATTATTATCGTGTTCGCCATTAGGGCAG :  46
CfMNPV  : TGGCATGTAGAACCTAATATCGTTATCTTGTTCACCATCGGGTTGG :  46
OpMNPV  : AGGCATGTAAAAGGTTACATCATTATCTTGTTCGCCATCCGGTTGG :  46
LdMNPV  : -GGAACGCTGATGGTCAATATTTTATCTAATCTGTTGCGCGGATAA :  45
TnSNPV  : TATGACGCTAAATGTAATTAGGCTATCTT-TATCTTGAG--GATAG :  43
SeMNPV  : TAACATACTAAACGTTACAATATTAGTAT-TATCTTGCGC--GTAG :  43
McNPV   : CAACATACTAAACGTGATAATATTGTCGT-TGTCTTTGGC--GCAG :  43
HzSNPV  : CACTATCGTAAACACAATTAAATTATCGATATGCGATAACGGTTGA :  46
HaSNPV  : CACTATCGTAAACACAATTAAATTATCGATATGCGATAACGGTTGA :  46
SpltNPV : ---CAACCAGCAACTTGTTCACGAACAGCAGCAACCGGTCTCCGAC :  43

AcMNPV  : TATAAATTGACGTTCATGTTGGATATTGTTTCAGTTGCAAGTTGAC :  91
BmNPV   : TATAAATTGACGTTCATGTTGAATATTGTTTCAGTTGCAAGTTGAC :  92
CfMNPV  : TATAAATAGACGTTCATGTTGGTTTTTGTTTCAGTTGCAAGTTGAC :  92
OpMNPV  : TATAAATAGACGTTCATGTTGGTTTTTGTTTCAGTTGCAAGTTGGC :  92
LdMNPV  : TATAAATTGATATTCATTTTTC---TTGTTTCAGTTGATCGTCGGC :  88
TnSNPV  : TATAAATTGACAATCATTTTATTTTTACTTCCAGTTC-TCACTTAC :  88
SeMNPV  : TATAAATTGACGTTCATGTTTGTAATTGTTTCAGTTT-GCTGAGAC :  88
McNPV   : TATAAATTGACGTTCATTTTGACGATTGTTTCAGTCT-GCCCGGTC :  88
HzSNPV  : TATAAATTGATGTTCATTTTTGTTTCAGAATTTATTG-AAATTGAA :  91
HaSNPV  : TATAAATTGATGTTCATTTTTGTTTCAGAATTTATTG-AAATTAAA :  91
SpltNPV : TATAACAGTTACAACAATTACAATAC-----CAGTCACCAAACTCC :  84
          TATA                           CAGT
AcMNPV  : ACTGGCGGCGACAAGATCGTGAACAACCAAGTGACTATG : 130
BmNPV   : ATTGGCGG-GACACGATCGTGAACAACCAAACGACTATG : 130
CfMNPV  : CGCGACGC-GCGCAGTACCTC-GCAATG----------- : 118
OpMNPV  : TGCGGCGC-GCGCAGCACCTTTGCGATG----------- : 119
LdMNPV  : GGGAA--ACGCCGAGTGGCAGGTGGACTCGCTTATG--- : 122
TnSNPV  : AA--GCGTTGTGATG------------------------ : 101
```


https://etd.uwc.ac.za/

```
SeMNPV  : AA--TCATG------------------------------  :  95
McNPV   : AC--ACGTCGTGAACATG---------------------  : 104
HzSNPV  : CTTTACGGCAAGTATG-----------------------  : 107
HaSNPV  : CTTTACGGCAAGTATG-----------------------  : 107
SpltNPV: AAGTGTCAAGAGGTCGAAGAAATTATG------------  : 111
```

Figure 2.3 Sequence alignment of *ie-1* promoter regions. Putative TATA boxes and consensus baculovirus early promoter elements (CAGT or CATT only for the HzSNPV and HaSNPV *ie-1s*) are indicated by either letters or in bold italic. Gaps are introduced with dashes for optimization of the alignment. Putative translation start codons (ATG) are in bold.

```
AcMNPV  : -------------MTQINFNAS-YTSASTPSRASFDNSYSEFCDKQ-  :  32
BmNPV   : -------------MTQINFNAS-YTSAPTPSRASFDNGYSEFCDKQQ  :  33
CfMNPV  : -----------MPKNMAALQQSLYTGPSTPSHTQFSR-STEFPENL-  :  34
OpMNPV  : -----------MPKNMETLQRS-YMGPSTPNHNLFNN-ATELPDDL-  :  33
McNPV   : ----------MHQAYKNCEGRSTPARDKLGSFLTFAPSTPEVIDQVF  :  37
SeMNPV  : -MHTPSHHNQISPATYKTTG-STPLRDSLGPFLKFQHSIPQPPLDMF  :  45
HzSNPV  : --MANRITTPLRDQVGNQVTINYPFQSQESCNYNNDSDSYMNRNNDV  :  45
HaSNPV  : --MANRITTPLRDQVGNQVTINYPFQSQESCNYNNDSDSYMNRNNDV  :  45
TnSNPV  : MSHPNIISAMDNNDASIKYKNYIDNAINTPTHTILQNVSMDFDDSNI  :  47
LdMNPV  : -----------------------------------------------  :   -
SpltNPV: ----------MDKLNDLNDHINIEKAFDESQYENFNSLFLEDQQEIM  :  37

AcMNPV  : PND----YLSYYNHPTPDGADTVISDSETAAASNFLASV--------  :  67
BmNPV   : PND----YLNYYNNPTPDGADTVVSDSRLQAASNFLASV--------  :  68
CfMNPV  : -------NFDVLNDSYETFSSVSLTT-----AE--------------  :  55
OpMNPV  : -------NFSTMDVPYDGSMPMNMS----------------------  :  51
McNPV   : G----------TDEPYTDNNYINLIENVNGINNTYD-----------  :  63
SeMNPV  : NNNNNNNTNQFVDIDYTDTNYNLLINNAEMINQNYDAAAAAAAASHH  :  92
HzSNPV  : DVKKLIKTVENASNKTVENASAFFASYIPPTSSNKPSPR--------  :  84
HaSNPV  : DVKKLIKTVENASNKTVENASAFFASYIPPTSSNKPSPR--------  :  84
TnSNPV  : LDFG--NENDMNVYDRRDSNSSKIVNDACDENSQFSDVNVNNNADNN  :  92
LdMNPV  : ------------MEPFLMYNEEMKTPAR-------------------  :  16
SpltNPV: PTPP--PSVREQPAPSVRKTPACDRPVLYGKGKRMKMSG--------  :  74

AcMNPV  : ----NSLT-DNDLVECLL------------KTTDN----LEEAVSSA  :  93
BmNPV   : ----NSLTDDNDIMECLL------------KTTDN----LGEAVSSA  :  95
CfMNPV  : ---------QDNQIDKIL------------QESAA----MNRDVNSE  :  77
OpMNPV  : ---------SDSLMNLLE------------DRSKK----LACAVDTE  :  73
McNPV   : ----DASNLQNINLNFFN-----------ALTSEN---FAPSMECNK  :  92
SeMNPV  : HDPFNASSPQDIDLDLLFPP--EQ--VPFVETKSNVDAVVPTIKIDK  : 135
HzSNPV  : --PNHLRFGDEIVMSPIAMS--PQRITPRSERSENVIESLPESLSSL  : 127
```


https://etd.uwc.ac.za/

```
HaSNPV  : --PNHLRFGDEIVMSPLAMS--PQRITPRSERSENVIESLPESLSSSL : 127
TnSNPV  : DYIKIMKTATDVVENKNEYTNKHKTAVVSTKPFKKNPKKRPSSSLTT : 139
LdMNPV  : ---------SDLQIMSAN------------EPEQTCLKLYSEFADFD :  42
SpltNPV: -----GKNMSKILNNNDIS-----------SSSSDDDDSGNETKSSS : 105

AcMNPV  : YYS----------------------ESLEQPVVEQP---------- : 107
BmNPV   : YNA----------------------ESFELPVAEQP---------- : 109
CfMNPV  : LA-----------------------QFTASEYVTG----------- :  89
OpMNPV  : LAR----------------------ESTASEFVAG----------- :  86
McNPV   : GNY----------------------VDIELP----------PKDE- : 105
SeMNPV  : KNMNIKDNNKENVAVVSPSLTLKHHSLSMSSPSSFTSKPIMSPLKKS : 182
HzSNPV  : KQVT---------------------VSLRRGSGLYG-------KNI : 145
HaSNPV  : KQVT---------------------VSLRRGSGLYG-------KNI : 145
TnSNPV  : TTTTTTEKKNKSRPNRPPNSTVIADGSIPPQPVIKPSKK---QTVFV : 183
LdMNPV  : EQY-----------------------AIEKEQYDER---------- :  55
SpltNPV: GGS-----------------------GFEKPKYKKK---------- : 118

AcMNPV  : S---PSSAYHAESFEHSAGVNQPSATGTKRKLDEYLDNS------QG : 145
BmNPV   : S---PSSAYNAESFEHPVGVNQPSATGTKRKLDEYLDDS------QS : 147
CfMNPV  : -----FRADTMEP---EVIVET-IGDSMKRKASELDSDSDS---GES : 124
OpMNPV  : -----FSADSPQAQLAETGAETGAAGGSKRKASEVDSDSDS---DDS : 125
McNPV   : --KPKNYIHVGKGAARKAVSMKDELKRNHSAIEEDDDDD-----DDD : 145
SeMNPV  : TIKEKRVSMKGMGAAMKRIRMTDTYDETMSSSSSSDDDDESGSSDDE : 229
HzSNPV  : QNLKENYEKTMDPYESDSSSLELTPKPKKRSNTEKKIAGVG--EKRS : 190
HaSNPV  : QNLKENYEKTMDPYESDSSSLELTPKPKKRSNTEKKIAGVG--KKRS : 190
TnSNPV  : SPLINRGGKNLNVLRNDNNNNFNNDSDDSNGSDSEDSDSTHPPPSKK : 230
LdMNPV  : -----QSAFQPQDYERARAVLEYQQALERGAEAESLVECAL---ADE :  94
SpltNPV: --KAAQPTSKPIISSSESSASSSDDEMAPSTVLPPVRDRSN---DDQ : 160

AcMNPV  : VVG--QFNKIKLRP-----------KYKKSTIQSCATLEQTINHN-- : 177
BmNPV   : VVG--QFNKNKLKP-----------KYKKSTIQSCATLEQTINHN-- : 179
CfMNPV  : SKGKKRVIKPKMRQ-----------RYKKATIQNKTSLTEECNYN-- : 158
OpMNPV  : SKGKKLVNKPKIRQ-----------RYKKATIQNRTSLTEERQYS-- : 159
McNPV   : DNGVNTTAKFFAS------------EGPQRTIIKQKTRGRYGKKM-- : 178
SeMNPV  : ANKIERNIKNFNDNDKNDDDIGAVVEDRERILIKPKSRGRYAKKMCV : 276
HzSNPV  : KKEKPATPLNEVG---------PVANMNKQLLMDDAPNRRYKQVHLK : 228
HaSNPV  : KKEKPATPLNEVGPVAN---IGPVANMNKQLLMDDAPNRRYKQVHLK : 234
TnSNPV  : TKMTSKSSKMSVTPQQQMPEILKINAADKNKVNDEKQTVKYNKKKQQ : 277
LdMNPV  : GAFVEKKKKKSKKR-----------EKRKLEVASSPPGKRMREDE-- : 128
SpltNPV: SEDRPDDDQSENRS-----------NDRSNDRPNDDQSENRSEDHR- : 195

AcMNPV  : -TNICTVASTQE-------ITHYFTND------------FAPYLMRF : 204
BmNPV   : -TNICTVASTQE-------ITHYFTND------------FAPYLMRF : 206
CfMNPV  : -TEICTVAPTDQ-------IAEYFKHD------------FSVYLEKQ : 185
OpMNPV  : -TEICTVAAPDQ-------IAKYFAQD------------FSAHLNEV : 186
McNPV   : -SALTKAHAQQ--------VQHDTATEQ----------LFDEILQDQ : 206
SeMNPV  : SSAMKPVHVETP-------TPSDPATDS----------LFREIITKQ : 306
HzSNPV  : PQHPQPRDPSNK-------CWP-IRFDE----------YMRTNVMPL : 257
HaSNPV  : PQHPQPRDPSEQ-------VLANPSLNE----------YMRTNVMPL : 264
TnSNPV  : SQDAGAVVVVKQQKLDNESTSQTSVNDDQQRSKDCDSPTNDLFENKI : 324
LdMNPV  : --DEPPPPPPPT-------MEIARQND------------FNTQIVYS : 154
SpltNPV: -SDEGRRTSFNG-------QQHSAAIN------------KIINYEDA : 222
```


https://etd.uwc.ac.za/

```
                      Ac213      I
AcMNPV  : D-----DNDYNSNRFSDHMSETGYYMFVVKKSEVKP----------- : 235
BmNPV   : D-----DNDYNSNRFSDHMSETGYYMFVVKKSEVKP----------- : 237
CfMNPV  : KS----DCQMSANRFSDYISETGYYVFVVKKSEHKP----------- : 217
OpMNPV  : KS----ECQMSANRFSDYISETGYYVFVVKKGDRKP----------- : 218
McNPV   : PS---DMHLQNNRLFASHLLDTGYYMFLVVYDQNSDDS--------- : 241
SeMNPV  : LQQNDDVVVDNSGIFTSHILDTSYYMFIISKSSNVDEV--------- : 344
HzSNPV  : VQNMPTFRVDKSRRFVDDIQQKNYHMFIVKEQENINSSSIE------ : 298
HaSNPV  : VQNMPTFRVDKSRRFVDRIQQKNYHMFIVKEQENINSSSIE------ : 305
TnSNPV  : IPNMMTMERDNNRKFVQYILNAHNYLFIVYENKYNAKTFNKNSNASI : 371
LdMNPV  : AEDARVSGGVRDKRFSTYVYNANYYMFIVSEDAETPARP-------- : 193
SpltNPV : DFS---AETDQSRRFVDFYTSKLYHMFIISPKTSDVDNDP-----AA : 261

                      Tn338          II
AcMNPV  : FEIIFAKYVSNVVYEYTNNYYMVDNRVFVVTFDKIRFMISYNLVKET : 282
BmNPV   : FEIIFAKYVSNVVYEYTNNYYMVDNRVFVVTFDKIRFMISYNLVKET : 284
CfMNPV  : FEVVFAKFVNNVTNEYTNNYYMVDNRVFVVSLNNVKFMVSYKLVREQ : 264
OpMNPV  : FEVVFAKFVNNATNEYTNNYYMVDNRVFVVSLNNVKFMVSYKLVREQ : 265
McNPV   : FVIRYVNCVHSVYNEYVARHMHHDRFVFVVTYERYRFMVSYKLLLHL : 288
SeMNPV  : YSLRYINCVHSVHNEYTAHHMHHDRFVLVVTIERYRFMISYNLLLGM : 391
HzSNPV  : HVILYANTVASINYEYSSYYYNVDKLVHVVTFNRYRFMISHFLLTKL : 345
HaSNPV  : HVILYANTVASINYEYSSYYYNVDKLVHVVTFNRYRFMISHRLLTKL : 352
TnSNPV  : YKIEYVNCVQSIYKYYNANYSHIDRTCKVVSFNRFRFAISVNLINKM : 418
LdMNPV  : FKVTYANCVWCINQDYRRNYRHVDNKVMVLSIDKCRFMISYKLLKKM : 240
SpltNPV : YELRYVNTVHSVLSEYRKYFSKLSNKVLVVTMARYRFMIVERVLSAM : 308

AcMNPV  : GIEIPHSQDVCNGETAAQNC-----KKCHFVDVHH-TFKAALTSYFN : 323
BmNPV   : GIEIPHSQDVCNDETAAQNC-----KKCHFVDVHH-TFKAALTSYFN : 325
CfMNPV  : GIDIPPHVNLCDDAQAERNP-----YDCYFEPVKN-VFQTTLINHFH : 305
OpMNPV  : GIDIPPHVNLCNDAQAERTP-----LNCYFEPVKN-AFQATLINHFH : 306
McNPV   : EIDIPQQDQFSETQLKNTNP-----KECYFEEVKNFEFLTFLTNYFH : 330
SeMNPV  : NIEIPTQDQFSEKQLSDTNK-----NMCIFEEVKDFKFLSLLINTFR : 433
HzSNPV  : NVHIPESEQFPMRVHQDRS------TKCHFNEIKDYVFMNELNHMFN : 386
HaSNPV  : NVHIPESEQFPMRVHQDAS------TKCHFNEIKDYVFMNELNHMFN : 393
TnSNPV  : QIELPPTEQFKKEDLKKISPKN---TFCLLNEVKDPDFISKLTNTFG : 462
LdMNPV  : NIPIPPSEDIERQAAAEAAEEAAREDKCYFNEIKDFEFLTLLINTFN : 287
SpltNPV : NITIPLLERIEDPKDNEIS----------FNEVKDSNEFNLLVHTFN : 345

                                  III
AcMNPV  : LDMYYAQTTFVTLLQSLGERKCGFLLSKLYEMYQDKNLFILPIMLSR : 370
BmNPV   : LDMYYAQTTFVTLLQSLGERKCGFLLSKLYEMYQDKNLFILPIMLSR : 372
CfMNPV  : LDMYYSQTTFVTLMQSMGESKSGMLLNKLYQMFQDRSLFTLPIMLSR : 352
OpMNPV  : LDMFYAQTTFVTLMQAVGENKTNMLLNKLYQMYQDRSLFTLPIMLSR : 353
McNPV   : LDKVYAQGKISLLLASIGEYKARLIYNTLTEMINDKSLFRLPFHMCK : 377
SeMNPV  : LDQVYIQGKVSLLLASVGESKSRVIFDQLTQMMDTGMMFILPMSVTK : 480
HzSNPV  : LDMVMVQTELYFLMSAIGPDKGKVLIKSVMEHINDDHLFVLPINLSR : 433
HaSNPV  : LDMVMVQTELYFLMSAIGPDKGKVLIKSVMEHINDDHLFVLPINLSR : 440
TnSNPV  : LDNIYIQGQLTMLLSSIGENRAKILNQHISAMIEDKSLFIIPLHLSR : 509
LdMNPV  : LDMCYTRVKIFMLLSSMGDSKSKMLWNWVYRVIKDETLFHIPVNYGH : 334
SpltNPV : LNTTIVQTDIAFMYSALSQSKAMYVHNKMNKLVADKTLFILPINVSR : 392

AcMNPV  : KESNEIE--------TASNNFFVSPYVSQILKYSES---VQ------ : 400
BmNPV   : KESNEIE--------TASNNFFVSPYVSQILKYSESIRKVK------ : 405
```


https://etd.uwc.ac.za/

```
CfMNPV : KEPTIEN--------TPLSRNYTSSYVAQIIKYSKN---VR------ : 382
OpMNPV : KEPVNEN--------APQNKNHAFSYVAQIMKYSKN---LR------ : 383
McNPV  : KEITD----------DEMSRAYSSAYVSDIIKLTEN---IK------ : 405
SeMNPV : KEAPNQ---------DELKKYDMSMYVEDIMKYTTG---LH------ : 509
HzSNPV : QESKLED--------IQRTVASVSLYVQNIVSLSKD---VQ------ : 463
HaSNPV : QESKLED--------IQRTVASVSLYVQNIVSLSKD---VQ------ : 470
TnSNPV : SKELEEIVDDDL---NPNNSNVSSAYIRDIIELSNK----------- : 542
LdMNPV : RQPIVEED-------FLAAPAAPECGASASSADSEHVKSVV------ : 368
SpltNPV: RDVVLEASEIAANVATPTSRAIAIATIESATNMSTKTPMIASDEQPL : 439

AcMNPV : ---------------FPDNPPNKYVVDNLNLIVNKKSTLTYKYSS-- : 430
BmNPV  : ---------------FPDNPPNKYVVDNLNLIVNKKSTLTYKYSS-- : 435
CfMNPV : ---------------FPENNPDNGVISRLEEIVTQKSSLTYKYSS-- : 412
OpMNPV : ---------------FPQGDPTQQVMDRLEEIVTQKSSLTYKYSS-- : 413
McNPV  : ---------------FKVLQEYKKKHDRPRVIKNVLNALSFWLRSKD : 437
SeMNPV : ---------------FNKFEE-DRKLSRAQIVDSVSKSLSFWYENKQ : 540
HzSNPV : ---------------FKQTAE--NFMNRDDVINYVTVALKFWLRSKN : 493
HaSNPV : ---------------FKQTAE--NFMNREDVINYVTVALKFWLRSKN : 500
TnSNPV : ---------------LKFKAPIIPSYVHKTKEQNIENVLSFWINIQK : 574
LdMNPV : ---------------SAGEGLSFRVADAKLTAEQALDSVRFWLRIKS : 400
SpltNPV: YISDIVKACHCASFVVNDHGAKSSPNKLESYIKKLTEELKFWLPNAS : 486

AcMNPV : --------------VANLLFNNYKYHDNIASNNNAENLKKVKKE-DG : 462
BmNPV  : --------------VANLLFNNYKYHDNIASNNNAENLKKVKKE-DG : 467
CfMNPV : --------------VANLLFSRYGHQ----RDNNADSLKKVKKE-DG : 440
OpMNPV : --------------VANLLFNRYGR-----RDNNADALKKVKKE-DG : 440
McNPV  : TKSSEFK----EKNNFTYKFGSVVRVLYDATDKNVSKLFKIKKE-NG : 479
SeMNPV : TIKNRNKQQQQEKSNFTYKYGCIARQFYDPTHKGVKKLFKVKKE-NG : 586
HzSNPV : EKNVVKE----QSDFFTYKYGSVVRLLEK-ESIHTNALLKIKRE-TG : 534
HaSNPV : EKNVVKE----QSDFFTYKYGSVVRLLEK-ESIHTNALLKIKRE-TG : 541
TnSNPV : NNNERDK-TLAKSLQFTYKFTSVARVLFDETDGDVNKLFKVKKE-PG : 619
LdMNPV : NDAVQKT----KDCYINYKYACIVRLLYDEQDKRIANLLKIKKPGAG : 443
SpltNPV: GEDVKKIR-----NSFTYKYSSVARLFYK--DRDLYLFKKVRKT-KG : 525
                                  IV
AcMNPV : SMHIVEQYLTQNVDNV-KGHNFIVLSFK--NEERLTIAKKNKEFYWI : 506
BmNPV  : SMHIVEQYLTQNVDNV-KGHNFIVLSFK--NEERLTIAKKNEEFYWI : 511
CfMNPV : NRLLVEQYMSQNENDE-TSHNFIVIQFGGVNDERLTIAKKGIEFFWI : 486
OpMNPV : NRLLVEQYMSYNENDD-TSHNFIVIQFGGVNDERLTIAKRGKEFYWI : 486
McNPV  : SVKLIEEYLSACKQFP-ESHNFILITTK--SDERITAIK------WA : 517
SeMNPV : STKLIENYLNACKERF-ENYSFILITTK--SDERITIIKNGMEFLWI : 630
HzSNPV : HAGLIDNYLEANQNDT-TSNSFILINTK--MDERITIIKKGPIFLWI : 578
HaSNPV : HAGLIDNYLEANQNDT-TSNSFILINTK--MDERITIIKKGPVFLWI : 585
TnSNPV : SVAMIEDYLQACEKIP-NGNNFIMINTL--NDERVTIIKAKNEFFWI : 663
LdMNPV : TAELVEHYLNVCAKLPRDSQNFLLVTTK--NEERLTLVKNGPRLVWI : 488
SpltNPV: NKSLVEAYLNAMNDHS--VHSFILVDTK--NEERLAIIKKNEVFVWI : 568

AcMNPV : SGEIKDVDV--SQVIQKYNRFKHHMFVIGKVNRRESTTLHNNLLKLL : 551
BmNPV  : SGEIKDVDA--SQVIQKYNRFKHHMFVISKVNRRESTTLHNNLLKLL : 556
CfMNPV : AAEIKDINV--DDLVKKYTRNVHHVFRIINVNRRESTTWHNNLLKLL : 531
OpMNPV : AGEIKDISV--DDLIKKYARNVHHVFRIINVNRRESTTWHNNLLKLL : 531
McNPV  : LSLSGSPPS-------------------------------------- : 526
SeMNPV : TSVIKDIIV--TDIIKKYKMYNHYVYNLNNGNRKEINIRHNGMIKLL : 675
```


https://etd.uwc.ac.za/

```
HzSNPV : TSIIKDIIA--MDLIEKYKKHTHAVFNLSNTNRKEMNNKHNGMIKLL : 623
HaSNPV : TSIIKDIIA--MDLIEKYKKHTHAVFNLSNTNRKEMNNKHNGMIKLL : 630
TnSNPV : RTNNPNNLIHCIDIIMAFKNFNHALLSLIPSNRKDLNNRHSGLIKLV : 710
LdMNPV : SGVARDICVG-DIINKFDGQFEHAVFKLNKVSRKELNNRHNGLLKLV : 534
SpltNPV: NCIMSSDIVP-EDIIHKHKDGTHYVFAMKRTNRKEVHARLNGMMKLI : 614
                          521

AcMNPV : ALILQGLVPLSDAITFAEQKLN--CKYKKFEFN-------------- : 582
BmNPV  : ALILQGLVPLSDAITFAEQKLN--CKYKKFEFN-------------- : 587
CfMNPV : QLLLQNLIRIDDVQQYSNKGDSK-FIYKRL----------------- : 560
OpMNPV : QLLLQNLIRLEDVQRYSDKSDTK-FVYKKV----------------- : 560
McNPV  : ----------------------------------------------- :   -
SeMNPV : SNYTGGRLTLNEATGIAVESFN--CNFEKVIYDKKNAKSIN------ : 714
HzSNPV : SFYTSNLLMLDELKEFAVNNFN--CSYDCKHYA-------------- : 654
HaSNPV : SFYTSNLLMLDELKEFAVNNFN--CSYDCKHYA-------------- : 661
TnSNPV : AYHLGGDVDINFVRAMAEKFKCN-YLYKKF----------------- : 739
LdMNPV : SLYTSAAVDLSVLVEIAQTQFE--CDYRCSQTSM------------- : 566
SpltNPV: STYVDESLKMKHVIKIAQDIFAANCEIIKYDDGRLVDRRDSAAVAAA : 661
                    570

AcMNPV : ---------------------------- :   -
BmNPV  : ---------------------------- :   -
CfMNPV : ---------------------------- :   -
OpMNPV : ---------------------------- :   -
McNPV  : ---------------------------- :   -
SeMNPV : ---------------------------- :   -
HzSNPV : ---------------------------- :   -
HaSNPV : ---------------------------- :   -
TnSNPV : ---------------------------- :   -
LdMNPV : ---------------------------- :   -
SpltNPV: VQQVKPAAIKRSVDDSSSIKKSKKQKV  : 688
```

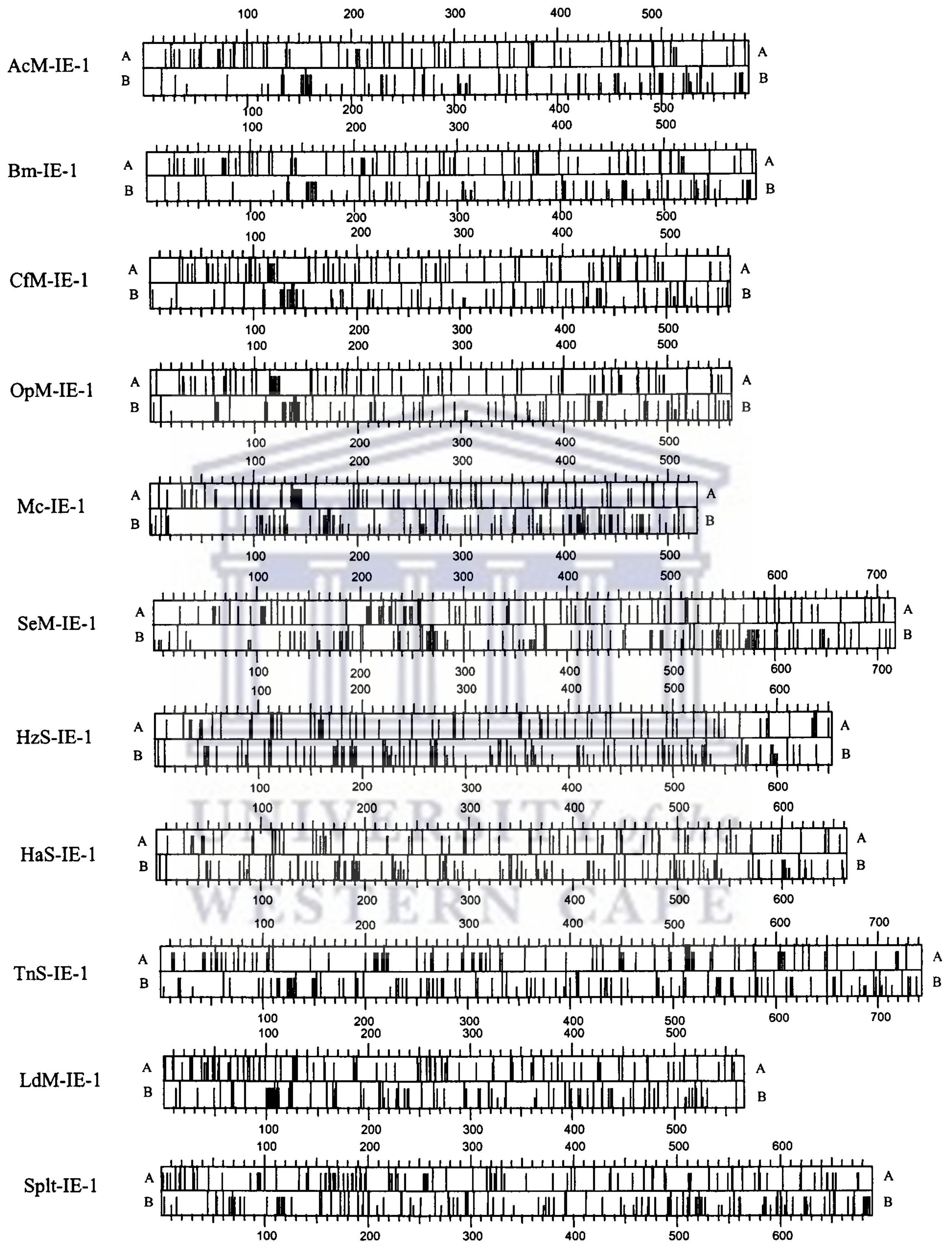
Figure 2.4 Alignment of amino acid sequences deduced from nucleotide sequences of eleven NPV IE-1s. The comparatively conserved regions of IE-1s start from the amino acids 213 (AcMNPV) and 338 (TnSNPV). The four conserved regions at C-termini are numbered in Roman. The position of the helix-loop-helix-like domain was indicated by numbers and arrows. The order of alignment sequences with GenBank accession number is: *Autographa califonica* (Ac) MNPV (P11138); *Bombyx mori* (Bm) NPV (P33245); *Choristoneura fumiferana* (Cf) MNPV (L04945); *Orgyia pseudotsugata* (Op) MNPV (P22114); *Mamestra configurata* (Mc) NPV(U59461); *Spodoptera exigua* MNPV (AF169823); *Helicoverpa zea* (Hz) SNPV (U67264); *Helicoverpa armigera* (Ha) SNPV

https://etd.uwc.ac.za/

(NC003094); *Trichoplusia ni* (Tn) SNPV (AF317709); *Lymantria dispar* (Ld) MNPV (AF006656); *Spodoptera litura* (Splt) NPV (NC003102).

A comparison of eleven deduced amino acid sequences of NPV IE-1s (Fig. 2.4) showed little similarity among the N-terminal regions. The acid-base maps of IE-1s (Fig. 2.5) showed that each of N-terminal regions contains a high proportion of acidic residues. The number of acidic residues in each of these regions was different from one another. A cluster of basic amino acids appeared to form the boundary of each acidic region, but the pattern and position of the clusters were different in each case. The comparison of charge distribution of N-termini of IE-s showed that the net negative charge remained in each case (Table 2.1). Two regions with net negative charge, designated as region I and II, were identified at the N-terminus of each NPV IE-1 (Table 2.1). The number of amino acid residues in each of the region Is was more than that of the corresponding region II, but the number of residues in the region II of the SpltNPV IE-1 was more than that of its region I. The net negative charge among the region Is ranged from -3 to -19, while among the region IIs from -3 to -15. SwissProt database search for all the full length IE-1 proteins revealed that the amino acids 47 to 120 aa of the deduced TnSNPV IE-1 protein significantly matched the SET-domain of transcriptional regulators, while no matches to the SET-domain were found in each of the other IE-1s.

https://etd.uwc.ac.za/



https://etd.uwc.ac.za/

Figure 2.5 Acid-base profiles of the eleven IE-1 proteins. The acidic and basic amino acids were indicated in lane A and lane B of each diagram, respectively. Medium and full bars in lane A represent aspartic and glutamic acids. Small, medium and full bars in lane B depict histidine, lysine and arginine.

	AcMNPV 582aa	BmNPV 587aa	CfMNPV 560aa	OpMNPV 560aa	McNPV 526aa	SeMNPV 714aa	HzSNPV 654aa	HaSNPV 661aa	TnSNPV 739aa	LdMNPV 566aa	SpltNPV 688aa
Region I	1-151	1-153	1-125	1-126	1-146	1-120	1-47	1-47	1-113	1-99	1-46
Segment size	(151aa)	(1-153)	(125aa)	(126aa)	(146aa)	(120aa)	(47aa)	(47aa)	(113aa)	(99aa)	(46aa)
Net charge	-16	-15	-14	-13	-15	-10	-3	-3	-13	-16	-8
Region II	161-228	163-230	148-183	143-211	190-230	199-262	149-170	149-170	192-228	122-166	139-230
Segment size	(67aa)	(67aa)	(35aa)	(68aa)	(40aa)	(63aa)	(21aa)	(21aa)	(36aa)	(44aa)	(91aa)
Net charge	-5	-5	-5	-3	-9	-11	-4	-5	-6	-4	-15

Table 2.1 The N-terminal net negative charge regions of eleven baculovirus IE-1s. The total amino acids of each deduced protein are shown under the corresponding IE-1. The position of each segment is indicated with numbers. The sizes of two amino acid segments of each IE-1 are indicated with numbers in brackets. Negative charge (-) in each case refers to net charge of the segment.

In contrast, the C-terminal regions were comparatively conserved, which contained four well conserved regions (Fig. 2.4). Apart from region I with 16 amino acids, the rest of three regions consisted of over 50 amino acids and most residues in each region were well conserved. More than half of the previously suggested helix-loop-helix-like domain was also conserved among the compared IE-1s. However, SwissProt database search showed that either these conserved regions, or the entire C-termini of the IE-1s, matched no homologies with relevant functional domains.

https://etd.uwc.ac.za/

2.4.2 Comparison and phylogenetic analysis of IE-1 proteins

Previously, AcMNPV, BmNPV, CfMNPV and OpMNPV were identified as group I NPVs and HzSNPV, HaSNPV, McNPV, SeMNPV, TnSNPV, LdMNPV and SpltNPV defined as group II NPVs (Cowan *et al.*, 1994; Bulach *et al.*, 1999). The alignment of deduced IE-1 amino acid sequences showed high degree of homologies among the AcMNPV, BmNPV, CfMNPV and OpMNPV IE-1s, whereas the rest of the NPV IE-1s (McNPV, SeMNPV, TnSNPV, LdMNPV and SpltNPV) shared much lower homologies, except HzSNPV and HaSNPV which displayed a higher degree of amino acid homology (Table 2.2). Among the group I and II IE-1s exist lower homologies (Table 2.2). This is in good agreement with the previously defined baculovirus phylogeny in which group I and II NPVs and GVs are distinct from one another (Zanotto *et al.*, 1993). The proposed IE-1s of subgroup II-A, SeMNPV and McNPV, shared 35 % identity, while subgroup II-B, between SpltNPV and either HzSNPV or HaSNPV, shared 21% identity (Table 2.2).

	BmNPV (%)	CfMNPV (%)	OpMNPV (%)	McNPV (%)	SeMNPV (%)	HzSNPV (%)	HaSNPV (%)	TnSNPV (%)	LdMNPV (%)	SpltNPV (%)
AcMNPV	96	43	44	20	18	21	21	16	20	16
BmNPV		43	43	20	18	21	21	16	19	15
CfMNPV			73	20	19	20	20	16	21	14
OpMNPV				19	19	20	20	16	21	14
McNPV					**35**	19	19	19	17	15
SeMNPV						24	25	**22**	19	17
HzSNPV							96	20	21	21
HaSNPV								20	21	21
TnSNPV									17	14
LdMNPV										17

Table 2.2 Identity values of the amino acid sequences of eleven baculovirus IE-1s. The percentage of identity amino acids shared by relevant sequences are indicated by

https://etd.uwc.ac.za/

numbers. The scores shared between TnSNPV and SeMNPV, and between SeMNPV and McNPV are in bold.

The TnSNPV IE-1 shared the highest homology with the SeMNPV homologue (22% identity and 40% similarity) and lower with the rest of IE- 1s (Table 2.3a). The comparison between the deduced full length TnSNPV IE-1 and other IE-1 homologues is proportionally less conserved than that of the C-terminal two third portions of the TnSNPV IE-1 with the other corresponding counterparts (Table 2.3a,b). In the latter case the TnSNPV IE-1 shared the highest homologies with SeMNPV (29% identity and 52% similarity) homologue (Table 2.3b).

(a)

	AcMNPV (%)	BmNPV (%)	CfMNPV (%)	OpMNPV (%)	McNPV (%)	SeMNPV (%)	HzSNPV (%)	HaSNPV (%)	LdMNPV (%)	SpltNPV (%)
Identity	16	16	16	16	19	**22**	20	20	17	14
Similarity	31	31	32	31	33	**40**	38	39	31	32

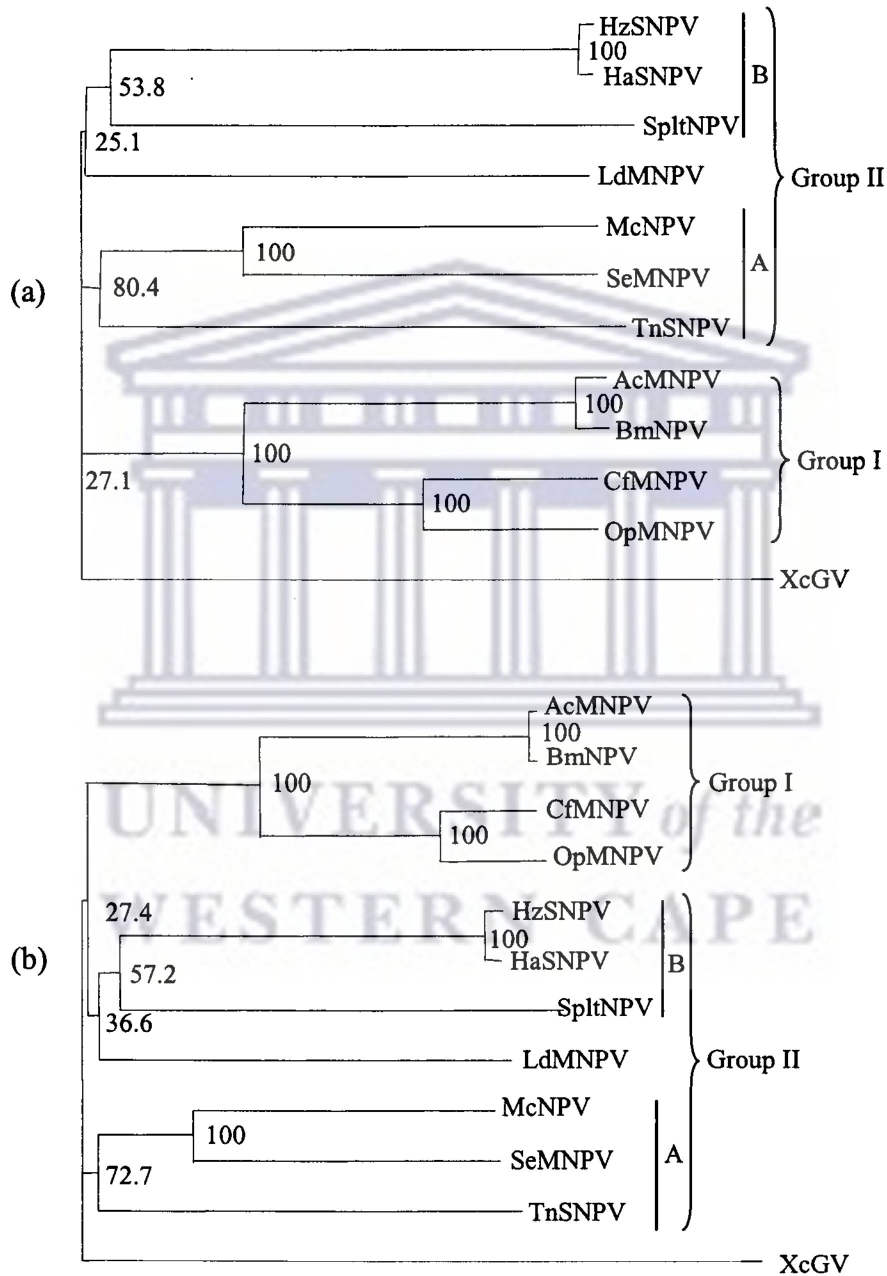
(b)

	AcMNPV (%)	BmNPV (%)	CfMNPV (%)	OpMNPV (%)	McNPV (%)	SeMNPV (%)	HzSNPV (%)	HaSNPV (%)	LdMNPV (%)	SpltNPV (%)
Identity	22	22	22	22	27	**29**	26	27	23	18
Similarit	39	39	39	39	44	**52**	46	47	41	35

Table 2.3 The full length (a) and the C-terminal two third portions (b) of ten deduced NPV IE-1s are compared with the TnSNPV homologue, respectively. The greatest homologies shared by TnSNPV and SeMNPV IE-1s in table a and b are in bold.

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

Based on alignment data sets of the predicted full length IE-1 proteins and the C-terminal two third portions of the amino acid sequences, two rooted phylogenetic trees, A and B, were constructed, respectively (Fig. 2.6a,b).



https://etd.uwc.ac.za/

Figure 2.6 Phylogenetic relationships of the predicted full length (a) and the C-terminal two thirds (b) of the IE-1 proteins. The bootstrap scores are present at each node as percentages of 1000 trials. Branch lengths are proportional to relatedness. The members of group I and II and proposed subgroup II-A and II-B are indicated in the trees. The XcGV taxa was used as an out-group.

The high bootstrap supports derived from trees were obtained for inferring phylogenetic relationships between the IE-1s of group I, subgroup II-A (SeMNPV and McNPV) and II-B (HzSNPV and HaSNPV) NPVs, but the separation of group I and II was supported by low bootstrap values in each tree, with a slightly higher score in tree B than in tree A. TnSNPV IE-1 obtained comparatively high scores (80.4% and 72.4 % for trees A and B) for being included in subgroup II-A, while the comparatively low bootstrap scores (53.8% and 57.25 % for trees A and B) were observed for inclusion of the SpltNPV in Subgroup II-B. Although the LdMNPV IE1 appeared to be clustered in subgroup II-B with the support of low bootstrap values (below 50%) in both trees, the bootstrap score of this branching in tree B was slightly higher than that in tree A. Similarly, the phylogenetic relationship between subgroup II-A and II-B NPV IE-1s was supported by low bootstrap values. In this case tree B showed a slightly higher bootstrap value than tree A in support of the separation.

2.5 Discussion

Many baculovirus early promoter regions contain a conserved early consensus motif

https://etd.uwc.ac.za/

CAGT and a TATA box (Friesen, 1997). To examine the functions and interactions of these early regulatory elements in directing transcription of the AcMNPV *ie-1*, a previous study (Pullen and Friesen, 1995b) demonstrated that site mutations of the promoter CAGT reduced the rate of transcription, but the RNA start site remained at or near the initiation site of the wild type *ie-1* for each mutation. In contrast, deletion of the TATA element reduced the rate of transcription much more than that of the former experiment but failed to abolish transcription from the promoter CAGT. However, deletion of the CAGT motif abolished early and late transcription (Pullen and Friesen, 1995a). This suggested that CAGT can function as a faithful initiator in a TATA-less promoter, while a TATA element in the *ie-1* promoter plays a role in stimulating transcription from the CAGT. Although Pullen and Friesen (1995b) demonstrated that the *ie-1* promoter CAGT motif and its flanking sequences were sufficient for proper transcription in a TATA-independent manner, the TATA plays an important role in cooperation with the CAGT initiator to properly regulate *ie-1* promoter activity. Since a TATA box is present in each of eleven NPV *ie-1* promoter regions, and an early promoter sequence CAGT highly conserved in the same region of most *ie-1*s (Fig. 2.3), both elements cooperating with each other may be a common feature for properly directing transcription of *ie-1*s. Although only a CATT motif was found in both HzSNPV and HaSNPV *ie-1* promoters, instead of a usual CAGT element, it may act as a transcription initiator since it conforms to the consensus early promoter CAG/TT (Friesen, 1997).

By observation of eleven NPV *ie-1* promoter regions, two CAGT motifs are found downstream of a putative TATA box in each case of CfMNPV and SpltNPV *ie-1*s,

https://etd.uwc.ac.za/

respectively. It has not yet been determined which CAGT motif could be utilized as RNA start site, but there is no available evidence that both CAGT motifs could act as initiators in the case of early transcription of *ie-1*. In addition, the study of Pullen and Friesen (1995b) suggested that the early promoter CAGT can function as a late promoter element in the case of the AcMNPV *ie-1* late transcription. The sequence gazing of these compared *ie-1* promoter regions did not find any late consensus promoter TAAG. Perhaps, as suggested earlier, a CAGT motif may play a role in late *ie-1* transcription.

Many transcriptional activator proteins contain an acidic domain as a common characteristic (Brent and Ptashne, 1985; Hope and Struhl, 1986; Trezenberg *et al*., 1988). The role of the N-terminal acidic region of the AcMNPV IE-1 as a transcriptional activation domain has been extensively studied in recent years. It was demonstrated that the N-termini of AcMNPV and OpMNPV IE-1s contain independent acidic transactivation domains, although sizes of the domains were mapped differently in the first 222 N-terminal amino acids (Kovacs *et al*1992; Rodems *et al*., 1997; Slack and Blissard, 1997; Forsythe *et al*., 1998;). Slack and Blisard (1997) identified two N-terminal acidic activation domains of the AcMNPV IE-1. One was mapped at amino acids 1-125, while the other at 168-222. Both domains contain net negative charge. These data are similar to those identified in this study. A comparison of eleven IE-1 amino acid sequences, including the AcMNPV IE-1 showed that the N-terminal residues are not conserved but rich in acidic features and dominant in negative charge. Although in most cases sizes of the acidic region Is are bigger than those of the region IIs, and the sizes between each of the region Is and each of the region IIs may be different from one

https://etd.uwc.ac.za/

another, the patterns of the net negative charge in the two identified acidic regions of each IE-1 are similar to those of AcMNPV IE-1 (Table 2.1). This appeared to suggest that IE-1 contains two acidic activation domains at its N-terminus as identified in the case of the AcMNPV and OpMNPV IE-1s (Slack and Blissard, 1997; Forsythe *et al.*, 1998). SwissPort database search revealed that the N-terminal 47-120 amino acids of the deduced TnSNPV IE-1 protein significantly matched the SET-domain of transcriptional regulators, although it was not the case for other IE-1s. The SET-domain has been found in more than 40 transcriptional regulation proteins from yeast to mammals. It is involved in modulating transcriptional activities as well as protein-protein interactions (Jenuwein *et al.*, 1998; Cui *et al.*, 1998). The position of the SET-like domain in the TnSNPV IE-1 is similar to that of one of the two acidic activation domains mapped previously in the case of the AcMNPV IE-1 (Slack and Blisard, 1997; Rodems *et al.*, 1997). This suggests that the SET-like domain of the TnSNPV IE-1 may serve as an acidic activation domain as did one of the AcMNPV IE-1 domains. In addition, it was reported that several basic amino acids were clustered together consisting of a boundary of the first N-terminal acidic region in the case of the OpMNPV and AcMNPV IE-1 (Theilmann and Stewart, 1990; Slack and Blisard, 1997). The clustered basic amino acid region was suggested to be a transcriptional inhibitory domain that may be negatively involved in ie-1 transcriptional regulation (Slack and Blisard, 1997). It appeared that all the IE-1s contain a cluster of basic amino acids at the border of its acidic region (Fig. 2.5). The proposed functional role of the clustered basic amino acids may be in common among IE-1s.

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

On the other hand, the most recent study (Pathakamuri and Theilmann, 2002) demonstrated that the N-terminal acidic activation domain contains an essential replication domain, but the functional roles of both domains are independent and separable from one another. This study suggested that since IE-1 binds to *hr*s, which are known as baculovirus replication origins, as a dimer, IE-1 may be involved in viral replication as an origin binding protein to interact with the other replication proteins to form the replication complex. However, it appeared that the acidic domain for replication could not be exchangeable for either AcMNPV or other heterologous counterparts. This conflicts with results from other studies which showed that native acidic activation domains for replication activation can be functionally replaced by heterologous counterparts (Cheng *et al.*, 1992; He *et al.*, 1993; Hu *et al.*, 1999; Li *et al.*, 1998). Thus, whether or not the independent acidic replication domain of the OpMNPV IE-1 is conserved in other IE-1s, or only specific to the OpMNPV IE-1, remains to be investigated.

Previously, the C-terminal half of the AcMNPV IE-1 has been identified as being involved in oligomerization and DNA binding (Rodems *et al.*, 1997 and Olsen *et al.*, 2001). Recently, van Strien *et al* (2000) compared the seven IE-1 amino acid sequences and found that the deduced proteins were comparatively conserved at their C-terminal two third regions. A comparison of eleven deduced baculovirus IE-1 protein sequences revealed that the sequence conservation at the C-terminal two thirds of the proteins mainly centered in four separate regions (Fig. 2.3). The region IV contains a helix-loop-helix-like domain that is involved in IE-1 oligomerization and DNA binding in the case

https://etd.uwc.ac.za/

of AcMNPV (Rodems *et al.*, 1997 and Olsen *et al.*, 2001). However, no significant match to any class of identified DNA binding domains were found within either four conserved regions, or entire C-termini, compared with entries in SwissProt database. Thus, apart from the helix-loop-helix-like domain, the functions of these conserved regions need to be further determined.

Previously, most analyses of baculovirus phylogenies were based on different genes such as *polyhedrin, dnapol, egt, gp41, chitinase, cathepsin, lef2, gp37, ie-1* and *vp39* (Bulach, *et al.*, 1999; Kang, *et al.*, 1998; Chen, *et al.*, 1999; Jin, *et al.*, 1999; Clarke, *et al.*, 1996; Li, *et al.*, 2001; Liu and Maruniak, 1999; van Strien *et al.*, 2000; Hayakawa *et al.*, 2000). However, these analyses have produced some conflicts between phylogenies based on different genes (Herniou *et al.*, 2003). In particular, *polh* phylogeny often disagrees with other gene phylogenies (Clarke, *et al.*, 1996; Kang, *et al.*, 1998). To resolve the conflicts among phylogenies of different genes, entire genomes have been used to reconstruct baculovirus phylogeny. Based on complete genome sequences several different approaches have been explored to infer baculovirus phylogenies, such as gene order, gene content, combined analysis of all shared genes in baculovirus genomes and comparison of complete genome sequences. With regard to the first two, each has some disadvantages for phylogenetic inference, such as lack of quantitative estimate and long branch attraction artifacts, while last two represent the most advanced and robust approaches to infer baculvirus phylogenies (Herniou *et al.*, 2001; 2003). However, the phylogenetic analysis of individual genes can still be used to bring some insight into the

https://etd.uwc.ac.za/

baculovirus phylogenies, as did the previous analyses using different genes for inference of baculovirus phylogenies, especially when there is a lack of entire genome information.

In previous attempts to study the TnSNPV phylogeny the virus polyhedrin protein and a 2.97kb region, containing *p26, p10*, AcMNPV ORF-29, *lef 6* and the C-terminal portion of *p74,* were employed separately to infer the phylogenetic relationship of TnSNPV among baculoviruses (Fielding and Davison, 1999; Fielding *et al.*, 2001). The following data sets were used in these studies: 1) sequence alignment of a single gene (*polh*) which provided limited phylogenetic information because of the small size of its amino acid sequence and high sequence conservation (Bulach *et al.*, 1999); 2) alignment comparisons of five deduced protein sequences (making up a total amino acids of 849 in size only) from TnSNPV; 3) the gene order of the five genes in a 2.97 kb fragment compared with the corresponding regions of four other NPVs. The studies suggested that TnSNPV is a member of group II NPV. The previous phylogenetic estimates were mainly based on comparing several deduced protein sequences, and were not supported by bootstrap analyses and, therefore, TnSNPV phylogenies could not be well resolved by such analyses, especially in the divergent group II NPVs. In contrast, *ie-1* gene, as one of six essential genes involved in viral DNA replication in AcMNPV and OpMNPV (McDougal and Guarino, 2000), is present in all the lepidopteran baculovirus genomes. The conservation of these essential genes in baculovirus genomes suggested that a common mechanism of the viral DNA replication cannot be substituted by host genes (Hayakawa *et al.*, 2000) and, therefore, the genes can be preserved during evolution. By contrast, some genes, such as *p26*, early *gp64* gene and *p10,* are not conserved in all the

https://etd.uwc.ac.za/

baculovirus genomes (Herniou *et al.*, 2001; Luque, *et al.*, 2001) and may not be suitable for phylogenetic studies because genes, such as the above, may not reflect whole baculovirus evolutionary history. On the other hand, the TnSNPV IE-1 has much longer amino acid sequence (739 aa) than that of *polh* (245-250 aa). Unlike *polh* with invariance of sequence in more than half among its homologues, the IE-1 contains both poorly conserved and comparatively better conserved amino acid sequences. In addition, the XcGV ie-1 gene provides an appropriate outgroup taxa since GVs are distinct from NPVs.

Recently, van Strien *et al.* (2000) conducted a phylogenetic analysis based on alignment of the C-terminal part of IE-1 proteins, but the bootstrap analysis of the full length proteins was not included in the study. To define the TnSNPV phylogeny in deep branch of group II baculoviruses, two data sets, based on alignments of the deduced full length IE-1s and the C-terminal two third conserved portions of the proteins, were combined to further explore the IE-1 phylogenies. The topologies of two trees constructed with the above two data sets were consistent with one another. Both trees each separated NPVs into groups I and II, and resolved group II NPVs into two subgroups (Fig. 2.6a,b). This is in agreement with previous baculovirus phylogenies based on two individual gene analyses (Bulach *et al.*, 1999). The phylogenetic relatedness between the TnSNPV IE-1 and its homologues (SeMNPV and McNPV IE-1s) of subgroup II-A was well supported by both trees with high bootstrap values and higher homologies compared between these IE-1s (Table 2.3a,b). Interestingly, it appeared that the weak bootstrap signals of branching in some areas of tree A were slightly reinforced in tree B (Fig. 2.6). This

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

implied that the poorly conserved amino acid sequences in the N-terminal first third portions of the full length proteins may produce a slightly negative impact on the confidence level of branching in some areas of tree A. However, the two phylogenetic estimates based on the two trees consistently supported the inference of TnSNPV phylogeny among baculoviruses. The analyses further confirm that TnSNPV is a member of group II NPVs and suggested that the virus is most closely related to the members of subgroup II-A SeMNPV and McNPV. In addition, although branching in the same areas of the trees for separation of group I and II and of subgroup II-A and II-B is weakly supported by the low bootstrap figures, the TnSNPV IE-1 protein appeared to share the most recent common ancestor with the subgroup II-A NPV homologues, and so did group I and II. Since relationships within group II NPVs are not well resolved (Herniou *et al.*, 2003), analysis to further resolve phylogenetic relatedness of baculovirus species in this group should be done when more ie-1 gene sequences become available. In addition, owing to a lack of the IE-1 sequence data from subgroup II-C defined earlier (Bulach *et al.*, 1999), what the relationship exists between TnSNPV and the members of group II-C, remains to be established.

UNIVERSITY of the
WESTERN CAPE

It is worth noting that, although the phylogenetic study of TnSNPV based on its IE-1 phylogeny suggested that TnSNPV may be a member of subgroup II-A, to accurately infer the virus phylogenies, analyses based on all combined genes shared between available baculovirus genomes, or comparison of complete genome sequences, need to be done.

https://etd.uwc.ac.za/

Chapter 3

UNIVERSITY of the
WESTERN CAPE

Experimental Prediction of a DNA-Protein Binding Region of the *Trichoplusia ni* Single Nucleocapsid Nucleopolyhedrvirus IE-1

https://etd.uwc.ac.za/

3.1 Abstract

The interaction of IE-1 (immediate early gene 1) with homologous region *(hr)* enhancer sequences is required for *hr* enhancer-dependent transactivation of IE-1 in baculovirus early gene regulation. To define which IE-1 region or domain directly binds to a *hr* enhancer, the *Autographa californica* (Ac) and *Trichoplusia ni* (Tn) nucleopolyhedrovirus (NPV) IE-1 deletion truncations fused with 6xHis-Tag of pET expression vector were expressed in the *E. coli* and purified. The IE-1 truncations contain a basic helix-loop-helix (b-HLH)-like domain. An electrophoretic mobility shift assay revealed that the AcMNPV IE-1 deletion truncation containing the C-terminal 130 amino acids can bind to a 60 bp repeat within *hr*5, directly and specifically. Due to the presence of highly conserved primary and secondary structures in the truncation regions of eight NPV IE-1s, the binding result suggested that the same protein-DNA binding activity may remain in the corresponding regions of the TnSNPV and other compared IE-1s. In addition, it was also suggested that it is likely that the b-HLH-like domain may be mainly responsible for the direct DNA binding. This study also provided protein sources for further Nuclear Magnetic Resonance (NMR) and crystallography assays.

https://etd.uwc.ac.za/

3.2 Introduction

The immediate early gene 1 (IE-1) is the principle transregulator for gene expression of *Autographa californica* nucleopolyhedrovirus (AcMNPV). Its transcriptional activation may occur in both sequence-independent (Blissard *et al.*, 1992) and dependent manners (Guarino and Summers, 1986b). The former requires only a basal promoter motif. In this case the specific DNA sequence appears not to be involved in recognition by IE-1. The latter is believed to be associated with enhancer elements for activation of promoters. IE-1 also negatively regulates transcription in the case of *pe38* and *ie-2* (Leisy *et al*, 1997). On the other hand, it was suggested that IE-1 may also contribute directly or indirectly to viral DNA replication (Kool *et al.*, 1993; Lu and Miller, 1995).

Homologous region (*hr*) enhancer sequences ranging in size from 30 to 800 bp are dispersed throughout the AcMNPV genome (Ayres *et al.*, 1994; Cochran and Faulkner, 1983). They may function not only as origin of viral DNA replication (Kool *et al.*, 1993; Leisy and Rohrmann, 1993) but also as enhancer elements for early gene expression (Rodems and Friesen, 1993 and 1995). IE-1-mediated transcriptional activation can be significantly amplified when the promoter is *cis* linked to homologous region (*hr*) enhancer elements from AcMNPV (Guarino and Summers 1986b; Guarino *et al*, 1986a; Pullen and Friesen 1995b), whereas interaction of IE-1with these sequences is essential for *hrs* to function as transcriptional enhancer (Leisy *et al*, 1995). Two previous studies were involved in identifying minimal sequences for DNA binding and enhancement of early gene expression. One suggested that a 24 bp palindrome within a 60 direct repeat

https://etd.uwc.ac.za/

(DR60) was the smallest DNA binding unit, although not functioning sufficiently as an enhancer element (Guarino and Dong, 1994). The other study indicated a 28-mer palindromic repeat was a minimal sequence for orientation- and position-independent enhancer activity (Rodems and Friesen 1995). Both palindromes are located in the left half of *hr5*. Current evidence shows that IE-1 binds to the 28 bp repeats as a dimer and both half-sites of the repeats are required for the optimal interaction of IE-1, but a single half of the repeats is not sufficient for the proper interaction of IE-1 with the 28-mer enhancer repeats (Rodems and Friesen 1995).

Many transcriptional activator proteins have two domains: one involved in regulation of other genes, and the other in the DNA-binding function (Ptashne, 1988; Mitchell and Tjian 1989; Brent and Ptashne 1985). The functional mapping studies of the AcMNPV IE-1 revealed that its N-terminus contains two transactivation domains, residue 8 to118 and 168 to 222 (Fig. 3.1A) , that are dispensable for DNA binding (Kovacs *et al.*, 1992; Rodems *et al.*, 1997; Slack and Blissard 1997), while the C-terminal half participates in DNA binding but not in transactivation (Fig. 3.1A) (Kovacs *et al.*, 1992; Rodems *et al.*, 1997). It was also reported that the basic-helix-loop-helix (b-HLH)-like motif at the C-terminus of IE-1 plays a critical role in the protein-DNA interaction and oligomerization that is required for IE-1 transactivation and contributes to the DNA-binding (Olson *et al.*, 2001; Rodems *et al.*, 1997). The basic-helix-loop-helix (b-HLH) domain was also identified in other transcriptional activators (Murre *et al.*, 1989). It plays the role in creating dimer combination by interaction between the amphipathic helices. This is an efficient mechanism for regulation of gene expression (Robinson and Lopes, 2000).

https://etd.uwc.ac.za/

In an attempt to define the role of the highly conserved C-terminal half region of the *Trichoplusia ni* single capsid nuclear polyhedrosis virus (TnSNPV) IE-1 in DNA-protein binding, and to provide information for studying the structure of the possible binding region by NMR or crystallography assay, N- and one half of C-terminal deletion truncations of the AcMNPV and TnSNPV IE-1s were constructed and expressed in the pET expression system (Novagen). The advantage of this approach was not only to eliminate the possible influences from the other potential binding residues or regions of the AcMNPV IE-1, but also to reduce the impact of possible cellular binding factors on DNA binding (Guarino and Dong, 1991) by protein purification. Unlike previous binding studies, the focus of this study was to identify an independent direct binding region other than determining which residues and regions contribute to DNA binding. However, due to lack of information of the TnSNPV *hrs*, which were thought to be a common feature in baculovirus genomes (Chen, *et al.*, 2000, 2001, 2002; Chapter 1), the AcMNPV IE-1 truncation was used as a model to experimentally predict the role of the corresponding region of the TnSNPV IE-1 in DNA-protein binding. In addition, the relationship between two independent enhancer elements (a 28-mer palindrome and a 60 bp direct repeat) was also discussed.

3.3 Materials and Methods

3.3.1 Prediction of the C-terminal secondary structures of NPV IE-1s

PSIPRED server was used to generate the predicted secondary structures of NPV IE-1s. The secondary structures of the relevant C-terminal regions were manually aligned

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

according to the corresponding amino acid sequence alignment of NPV IE-1s (Wang *et al.*, 2001) generated from ClustalW (Thompson *et al.*, 1994).

3.3.2 Plasmid constructs

The N- and C-terminal deletion mutants used in this study were generated by PCR using a AcMNPV-IE-1 plasmid kindly obtained from Dr. Knebel-Morsdorf (University of Cologne, Germany) and a TnSNPV IE-1 fragment (Wang *et al.*, 2001) as templates. The deletion-specific primers used in PCR were as follows: AcIE453/421F, 5'-cacacaca-ggatccAATTTAAAAAAGGTTAAGAAG GAG-3'; AcIE582R, 5'-tctctctcaagctt-TTAATTAAATTCGAATTTTTTATATTTAC-3'; TnIE647/310F, 5'- cacacacaggatcc-AACGATGAACGCGTGAC-3'; TnIE739R, 5'-ctctctctaagcttTTAAAATTTTTTAT-ATAGATAATTAC-3'. A *Bam*HI and a *Hind*III sites (underlined in the deletion-specific primers above) were introduced into the 5' and 3' ends of the mutants by PCR using *pfu* DNA polymerase (Promega), respectively. The PCR products were purified using a PCR purification kit (Qiagen) and digested with *Bam*HI and *Hind*III endonucleases according the manufacturers' instructions (Promega). The standard cloning techniques (Sambrook *et al.*, 1989) were used to insert the digested PCR products into the corresponding sites of the pET28a expression vector (Novagen). The resulting plasmids were confirmed to be in frame by sequencing (Wang *et al.*, 2001) using both T7 promoter and terminator primers. Predicted sizes of the IE-1 deletion truncations, AcΔ1-452 and TnΔ1-646, and their amino acid sequences, are shown in Fig. 1A and B.

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

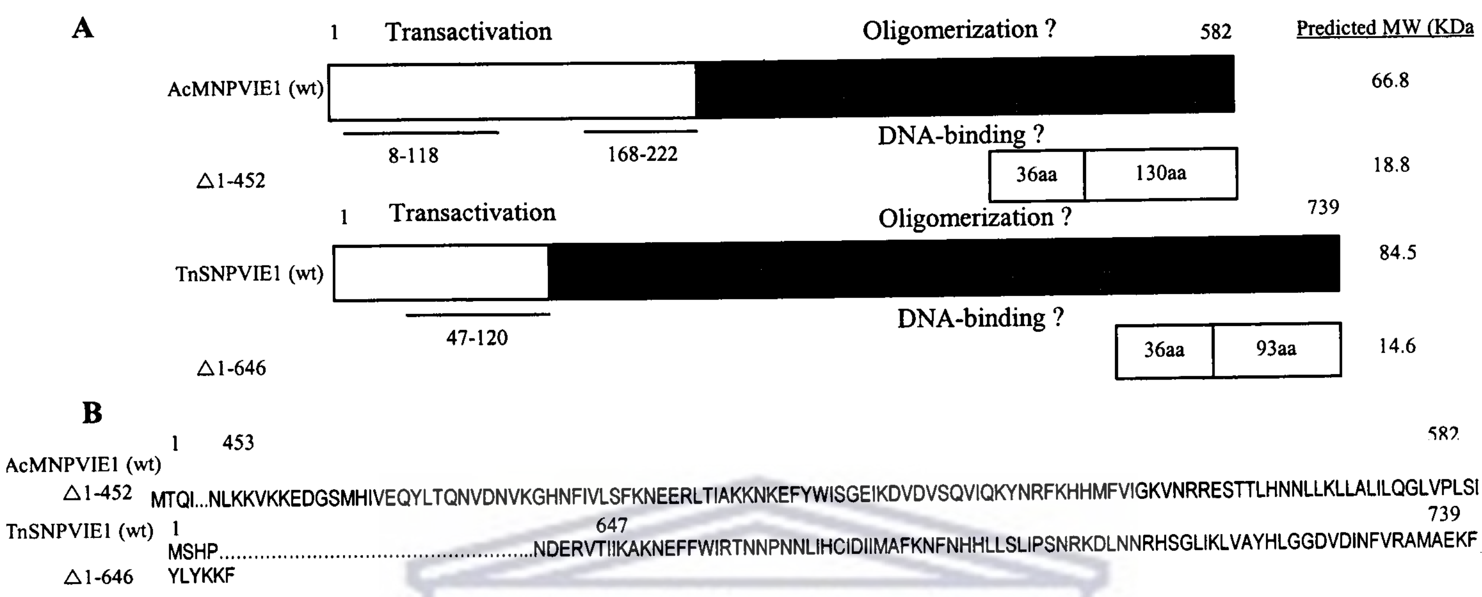


Figure 3.1 Schematic representation and amino acids sequences of the wild type IE-1s and their truncations. (A) Functional structures of the IE-1s and their 6xHis-tagged truncation fusions. Transactivation domains of AcMNPV (residue 8 to 118 and 168 to 222) and TnSNPV IE1s (residues 47 to 120) (see the text) are within their N-terminal half and unmapped DNA-binding and oligomerization domains within their C-terminal half. A N-terminal His-tag was fused with IE1 truncations. The sizes of the full length proteins and their truncation fusions in MW are indicated. The entire His-tag configuration and the two truncations have 36, 93 and 130 amino acids, respectively. The symbol Δ and following numbers denote the deleted amino acids from the IE-1s. (B) The amino acid sequences of the wild type IE-1s and their truncations. The different dashed lines represent the amino acids omitted and removed from the wild types and truncations, respectively. The positions of the amino acids are indicated in numbers.

https://etd.uwc.ac.za/

3.3.3 Protein expression and purification

All the procedures for the protein expression were followed according to the manufacturer's instruction (Invitrogen). Briefly, all the plasmids were transformed into *E.coli* BL21(DE3) pLysS strain. The cells were propagated at 37°C until OD_{600} reached 0.6. The following protein induction of cells was carried out at 18°C. A final concentration of 1mM IPTG was used for induction of the proteins. The samples were collected before and after 1h, 2h and 3h induction and analyzed by a 12% polyacrylamide gel according to a standard protocol (Sambrook *et al.*, 1989). The buffer components used for the protein purification under the denaturing condition were as follows: 1) Binding buffer: 8M urea, 20 mM sodium phosphate and 500 mM sodium chloride (pH to 7.8); 2) Wash buffer 1: 8M urea, 20 mM sodium phosphate and 500 mM sodium chloride (pH 6.0); 3) Wash buffer 2: 8M urea, 20 mM sodium phosphate and 500 mM sodium chloride (pH 5.3); 4) Elution buffer: 8M urea, 20 mM sodium phosphate and 500 mM sodium chloride (pH 4.0). Protein purification was carried out following the manufacturer's instruction (Invitrogen). Briefly, 4 ml of the resin was prepared by washing with H_2O three times. The 100 ml cell pellets for each sample collected after 3h induction were lysed in a 10 ml urea denaturing buffer and sonicated using six 10-second bursts at a medium intensity setting with a 10-second cooling period between each burst. The lysates were centrifuged at 3000 x g for 15 min at 4°C. The 5 ml supernatant aliquot of each sample was mixed with the 2ml prepared resin. The mixture was gently rocked for 30 min and then the resin was settled with low speed centrifugation (800 x g). The supernatant was carefully aspirated. The procedure was repeated with a second 5 ml aliquot. The lysate-resin mixture (supernatant) was loaded into an empty column,

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

followed by three washes. The His-tagged fusion proteins were eluted by applying 5 ml of denaturing elution buffer with pH 4.0. The purified protein fractions were confirmed by SDS-PAGE. The fractions were dialyzed and refolded using 3500 MW cut off dialysis tubing (Pierce) against 10 mM Tris, pH 8.0 with 0.1% Triton X-100 overnight at 4°C. The dialyzed protein fractions were stored at 4°C.

3.3.4 Western blot assay

The supernatants and precipitants of the purified protein fractions were separated by centrifugation (10000 x g) for 10 min and resolved onto a 15 % polyacrylamide gel. The cell lysates obtained from the *E. coli* strain, which contained pET vector only, were used as negative controls. The proteins on the gel were transferred onto a nitrocellulose membrane, blocked with blocking buffer, washed and detected by Ni-NTA AP conjugate according to the manufacturer's protocol (Qiagen). Ni-NTA conjugates consisted of Ni-NTA coupled to calf intestinal alkaline phosphatase. They can be used for direct detection of any fusion proteins with 6xHis tag. The detection was based on the affinity of nickel-nitrilotriacetic acid (Ni-NTA) metal-affinity chromatography matrices to 6 consecutive histidine residues which were tagged with the N-termini of the target proteins in this study. The reliability of the technology for 6xHis tagged protein detection has been recommended by the manufacturer's handbook (Qiagen). Although the positive control was not available, the results from the detection were positive and reliable, based on the sizes of the truncations, strong positive signals and the indications of the negative controls (the cell lysate obtained from the propagated *E. coli strain* containing a pET

https://etd.uwc.ac.za/

vector only) which were similar to the test samples but lacking the target proteins (AcΔ1-452 and TnΔ1-646).

3.3.5 Radio-labeling of the probe and electrophoretic mobility shift assay (EMSA)

The complementary sequences of a 60 bp repeat within the left half of the AcMNPV *hr5* were synthesized and annealed together. The SP1 21 bp oligo used for a positive control was provided from a gel shift kit (Promega). The two oligos (the 60 bp repeat and SP1 21 bp oligo) were end labeled with ^{32}P using T4 polynucleotide kinase and purified by G-25 spin columns according to the manufacturer's protocol (Roche). The gel shift assay was performed using the gel shift kit, following the protocol recommended by the manufacturer (Promega). The HeLa nuclear extract provided from the kit and the SP1 21 bp probe were used as a positive control in the experiment. A purified fish viral protein, cloned into the same sites of the pET28a vector as those of the truncation Δ1-452, was used as a negative control. 1μl ^{32}P labeled 60 bp probe (2 pmol/μl) was incubated with 3 μl of the fish viral protein. The purified AcMNPV IE-1 truncation (Δ1-452) was centrifuged at 10000 x g for 10 min at 4°C and the supernatant (soluble fraction) was used in EMSA. Due to the low concentration of the soluble fraction of the purified truncation Δ1-452, 6μl of the fraction and 1μl ^{32}P labeled 60 bp probe (2 pmol/μl) were mixed in the reaction. The amounts of the cold specific oligos used in the competition assay were 2 μl SP1 21 bp oligo (2 pmol/μl) for the positive control and 1.5 μl and 2 μl 60 bp oligo (2 pmol/μl) for Δ1-452, respectively. The 2 μl non-specific AP2 oligo (2 pmol/μl) was incubated with HeLa nuclear extract and the same amount of the cold non-specific SP1 oligo with Δ1-452. The control and Δ1-452 groups (Fig. 3.5A and B) were

https://etd.uwc.ac.za/

resolved in a 4 % and 6% polyacrylamide gel in 0.5x TBE buffer, respectively, and ran for 20 min at 350V. The gels were covered with plastic wrap and exposed to X-ray film overnight at -70°C.

3.4 Results

3.4.1 Prediction of the secondary structures and the experimental design

To identify the DNA-protein binding region of the TnSNPV IE-1, the Δ1-646 and Δ1-452 deletion mutants of the TnSNPV and AcMNPV IE-1s were constructed and expressed. Initially, one of the applications for the AcMNPV IE-1 truncation Δ1-452 was designed to be a positive control. However, due to lack of information of the TnSNPV *hr* repeats, which may bind to the Δ1-646 of the TnSNPV IE-1, prediction of the secondary structures of eight NPV IE-1s was conducted. It was found that the structures of half of the C-termini of IE-1s are highly conserved (Fig. 3.2A), especially in the previously suggested basic-helix-loop-helix (HLH)-like domain (Rodems *et al.*, 1997) in which the predicted secondary structures are almost completely conserved (Fig. 3.2A). Therefore, it was decided to employ the AcMNPV IE-1 truncation Δ1-452, instead of the Δ1-646 of the TnSNPV IE-1, to carry out the DNA-protein binding assay. The purpose of this was not only to identify whether the AcMNPV IE-1 truncation may directly bind to the AcMNPV enhancer element, but also to experimentally predict the corresponding binding region of the TnSNPV IE-1.

A

```
AcMNPV    -LLFNNYKYHDNI----A-SNNNAENLKKVKKE-DGSMHIVEQYLT  472
BmNPV     -LLFNNYKYHDNI----A-SNNNAENLKKVKKE-DGSMHIVEQYLT  477
```


https://etd.uwc.ac.za/

```
CfMNPV   -LLFSRYGHQ----------RDNNADSLKKVKKE-DGNRLLVEQYMS   450
OpMNPV   -LLFNRYGR-----------RDNNADALKKVKKE-DGNRLLVEQYMS   450
HzSNPV   -SDFFTYKYGSVVRLLFK-ESIHTNALLKIKRE-TGHAGLIDNYLE   544
SeMNPV   EKSNFTYKYGCIARQFYDPTHKGVKKLFKVKKE-NGSTKLIENYLN   596
LdMNPV   -DCYINYKYACIVRLLYDEQDKRIANLLKIKKPGAGTAELVEHYLN   453
TnSNPV   KSLQFTYKFTSVARVLFDETDGDVNKLFKVKKE-PGSVAMIEDYLQ   629

AcMNPV                       C----C-CCCCHHHHHHHCCC-CCCHHHHHHHHH
BmNPV                        C----C-CCCCHHHHHHHCCC-CCCHHHHHHHHH
CfMNPV                       -------CCCCHHHHHHHCCC-CCCHHHHHHHHH
OpMNPV                       -------CCCCHHHHHCCCCC-CCCHHHHHHHHH
HzSNPV                       EEEEEC-CCCCHHHHHHCCCC-CCCHHHHHHHHH
SeMNPV                       HHHHHCCCCCCHHHHHHHCCC-CCCHHHHHHHHH
LdMNPV                       EEEEECCCCCCHHHHHHCCCCCCCCHHHHHHHHH
TnSNPV                       HHHHHCCCCCCHHHHHHCCCC-CCCHHHHHHHHH

AcMNPV   QNVDN-VKGHNFIVLSFK--NEERLTIAKKNKEFYWISGEIKD--V   513
BmNPV    QNVDN-VKGHNFIVLSFK--NEERLTIAKKNEEFYWISGEIKD--V   518
CfMNPV   QNEND-ETSHNFIVLQFGGVNDERLTIAKKCIEFEWIAAEIKD--I   493
OpMNPV   YNEND-DTSHNFIVLQFGGVNDERLTIAKKCKEFYWIAGEIKD--I   493
HZSNPV   ANQND-TTSNSFILINTK--MDERITIIKKGPIFLWITSIIKD--I   585
SeMNPV   ACKER-FENYSFILITTK--SDERITIIKNGMEFLWITSVIKD--I   637
LdMPNV   VCAKLPRDSQNFLLVTTK--NEERLTLVKNGPRLVWISGVARD--I   495
TnsNPV   ACEKI-PNGNNFIMINTL--NDERVTIIKAKNEFEWIRTNNPNNLL   672

AcMNPV   HCCCC-CCCCCEEEEEEC--CCCEEEEEEECCCEEEEEEEHHHHH--C
BmNPV    HHHCC-CCCCEEEEEEECC--CCHHHHHHHCCCEEEEEEECCCCC--C
CfMNPV   HHCCC-CCCCEEEEEEECCCCCHHHHHHHHCCCEEEEEEECCCCC--C
OpMNPV   HHCCC-CCCCEEEEEEECCCCCHHHHHHHHCCCEEEEEEECCCCC--C
HzSNPV   HHCCC-CCCCEEEEEEECC--CCHHHHHHHCCCEEEEEEECCCCC--C
SeMNPV   HCCCC-CCCCCEEEEEEC--CCCEEEEEEECCCEEEEEEEEECCC--C
LdMNPV   HHHCCCCCCCCEEEEEECC--CCHHHHHHHCCCEEEEEEECCCCC--C
TnSNPV   HHCCC-CCCCEEEEEEECC--CCEEEEEEECCCEEEEEEEECCCCEEC

AcMNPV   DVSQVIQKYN-RFKHHMFVIGKVNRRESTTLHNNLLKLLALILQGL   558
BmNPV    DASQVIQKYN-RFKHHMFVISKVNRRESTTLHNNLLKLLALILQGL   563
CfMNPV   NVDDLVKKYT-RNVHHVFRIINVNRRESTTWHNNLLKLLQLLLQNL   538
OpMNPV   SVDDLIKKYA-RNVHHVFRIINVNRRESTTWHNNLLKLLQLLLQNL   538
HZSNPV   IAMDLIEKYK-KHTHHVFNLSNTNRKEMNNKHNGMIKLLSFYTSNL   630
SeMNPV   IVTDIIKKYK-MYNHYVYNLNNGNRKEINIRHNGMIKLLSNYTGGR   682
LdMPNV   CVGDIINKFDGQFEHHVFKLNKVSRKELNNRHNGLLKLVSLYTSAA   541
TnsNPV   HCIDIIMAFK-NFNHHLLSLIPSNRKDLNNRHSGLIKLVAYHLGGD   717
               521                     543

AcMNPV   CHHHHHHHHH-HCCEEEEEEEECCCHHHHHHHHCCHHHHHHHHHCCC
BmNPV    CHHHHHHHHH-CCCEEEEEEECCCCHHHHHHHHHHHHHHHHHHHHCCC
CfMNPV   CHHHHHHHHH-CCCEEEEEEECCCCHHHHHHHHHHHHHHHHHHHHCCC
OpMNPV   CHHHHHHHHH-CCCEEEEEEECCCCHHHHHHHHHHHHHHHHHHHHCCC
HzSNPV   CHHHHHHHHH-CCCEEEEEEECCCCHHHHHHHHHHHHHHHHHHHHCCC
SeMNPV   CHHHHHHHHH-HCCEEEEEEEECCCHHHHHHHHCCHHHHHHHHHCCC
LdMNPV   CHHHHHHHHCCCCEEEEEEEECCCCHHHHHHHHHHHHHHHHHHHHCCC
TnSNPV   CHHHHHHHHH-CCCEEEEEEEEECCHHHHHHHHHHHHHHHHHHHHCCC

AcMNPV   VPLSDAITDAEQKLN-CKYKKFEFN--------   582
BmNPV    VPLSDAITDAEQKLN-CKYKKFEFN--------   587
CfMNPV   IRIDDVQQYSNKGDSKFIYKRL-----------   560
```


https://etd.uwc.ac.za/

```
OpMNPV    IRLEDVQRYSDKSDTKFVYKKV-----------  560
HzSNPV    LMLDELKEFAVNNFN-CSYDCKHYA--------  654
SeMNPV    LTLNEATGIAVESFN-CNFEKVIYDKKNAKSIN  714
LdMPNV    VDLSVLVEIAQTQFE-CDYRCSQTSM-------  566
TnSNPV    VDINFVRAMAEKFKCNYLYKKF-----------  739
                    568
AcMNPV    CCHHHHHHHHHHHCC-CCCEEEECC--------
BmNPV     CCHHHHHHHHHHHCC-CCEEEEECC--------
CfMNPV    CCHHHHHHHHHHCCCCEEEECC-----------
OpMNPV    CCHHHHHHHHHHCCCCEEEECC-----------
HzSNPV    CCHHHHHHHHHHHCC-CCCEECCCC--------
SeMNPV    CCHHHHHHHHHHHCC-CCCCEEEECCCCCCCCC
LdMNPV    CCHHHHHHHHHHHHC-CCCCCCCCCC-------
TnSNPV    CCHHHHHHHHHHCCCCEEEECC-----------
```

B

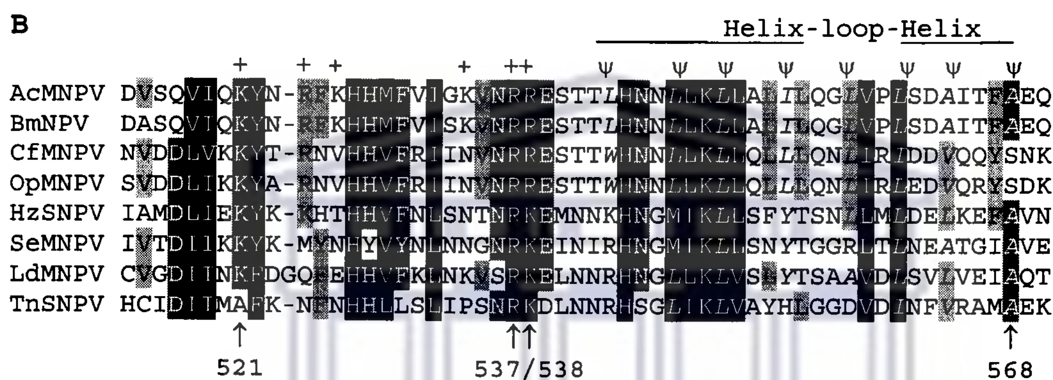


Figure 3.2 Structural alignments of the putative C-terminal binding regions and the b-HLH-like domains of eight IE-1s. (A) Alignment of the predicted secondary structures of half of the C-termini of the NPV IE1s. Each amino acid residue is aligned with its corresponding secondary structure. The gaps introduced in the alignment are copied as precisely as those in each aligned amino acid sequence. The helix structures of the basic helix-loop-helix-like domain are underlined and the position of the domain of the AcMNPV IE-1 is indicated in numbers. The letters, H, C and E in the predicted secondary structures represent helix, coil and strand, respectively. (B) Comparison of HLH-like domains in the same viral proteins. Eight hydrophobic (ψ) (Italic) and several conserved basic (+) residues are indicated (see text). The numbers represent the position of amino acids of the AcMNPV IE1 domain.

https://etd.uwc.ac.za/

3.4.2 Expression, purification and detection of the AcMNPV and TnSNPV IE1 truncations

The uninduced and induced samples (3h induction) of the truncations Δ1-452 and Δ1-646 were resolved on a 12% polyacrylamide gel (Fig. 3.3A). The evidence for expression was clearly exhibited on the gel. As expected, the sizes of the truncations were the same as predicted ones (Fig. 3.3A). The expression levels from the samples after 3h induction appeared not to be high. However, the purified fractions 3 and 4 of the Δ1-452 and 3,4, and 5 of Δ1-646 showed much higher intensity in the gel than those of induced samples (Fig. 3.3A, B), although only 6 μl of each fraction was loaded in the gel. The sizes of the purified truncations are the same as those of unpurified samples.

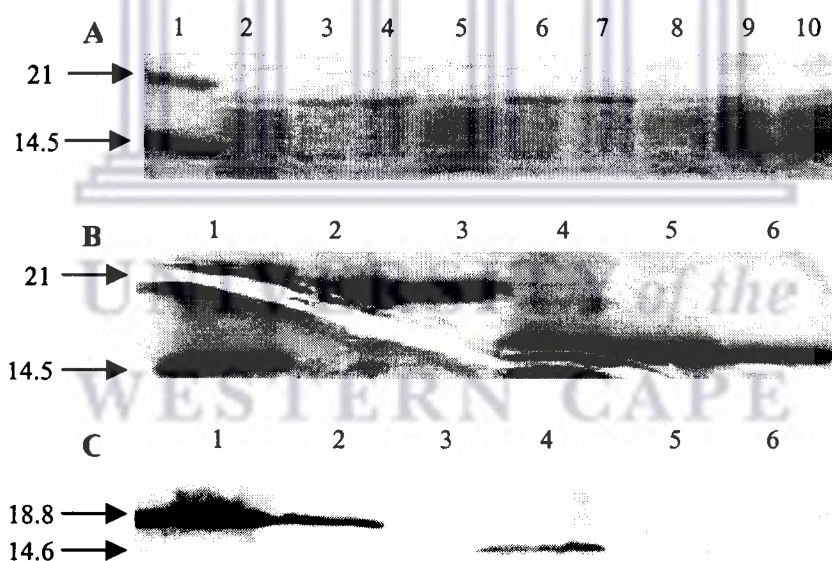
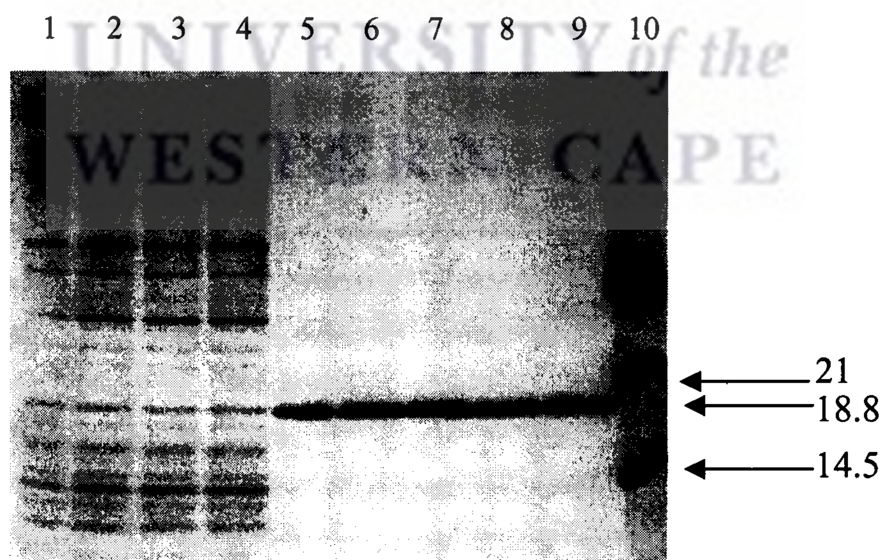


Figure 3.3 Expression, purification and detection of the AcMNPV and TnSNPV IE-1 truncations. (A) Expression of IE-1 Δ 1-452 and Δ 1-646. The uninduced sample of Δ 1-452 is shown in lane 2 and 5, while the uninduced sample of Δ 1-646 in lane 8. The

https://etd.uwc.ac.za/

sample of Δ 1-452 collected from 3h induction was loaded in lane 3, 4, 6 and 7 (duplicates of the same sample), while the samples of Δ 1-646 from 3h induction in lane 9 and 10 (duplicates). (B) Purified fractions of the IE-1 truncations. The fraction 3 and 4 of Δ 1-452 are in lane 2 and 3. The fraction 3, 4 and 5 of Δ 1-646 are indicated in lane 4, 5 and 6. (C) Detection of the two purified truncations. The precipitants of both truncations are detected in lane 1 (Δ 1-452) and 4 (Δ 1-646) and supernatants are in lane 2 (Δ 1-452) and 5 (Δ 1-646), respectively. Nothing is observed in the negative controls (lane 3 and 6). The sizes of the truncations (KDa) are shown on the left side of the Figure C, while the molecular sizes of the protein standard in KDa on the left side of Figure A and B.

Most of the proteins produced from *E.coli* were washed away through the purification column and a strong signal of the 6xHis tagged truncation (Δ1-452) was detected by Coomassie blue staining (Fig. 3.4).



https://etd.uwc.ac.za/

Figure 3.4 Efficacy of purification of the 6xHis tagged truncation (Δ1-452). The flow-through (non-6xHis tagged proteins) of the wash was loaded on lane 1-4. The purified truncation (Δ1-452) was eluted from the column in lane 5-9. The numbers represent the sizes of the protein ladder and the truncation in KDa.

Upon dialysis the supernatants and precipitants (aggregated protein fractions) of both purified samples were detected by western blot. Although soluble fractions of the truncations in the gel were hardly to be seen, as was observed by a low intensity of the Coomassie blue staining (data not shown), they were still detected in the western blot (Fig. 3.3C). In this experiment the concentrations of the proteins were not quantified but, according to the manufacturer's instruction (Promega), the two different amounts (1.5μl and 2μl) of the DNA probe with a low concentration (2 pmol/μl) were used for the DNA-protein reaction. A comparison of these two results could give some idea whether an excess of protein (if any) was in the binding reaction.

3.4.3 Gel shift assay

UNIVERSITY of the WESTERN CAPE

An EMSA was first performed on the control groups to verify whether the controls and the system can work as expected. A fish viral protein carried by pET28a vector was incubated with the ^{32}P labeled 60 bp repeat which was previously identified within the left half of AcMNPV *hr*5 (Guarino and Dong, 1991). In this case there was no indication showing that a DNA-protein complex was formed on the gel (Fig. 3.5A). This indicated that the 60 bp repeat does not respond to the vector-encoded N-terminal fusion and the fish viral protein. In the positive control a transcription factor protein in the HeLa nuclear extract interacted with the ^{32}P labeled SP1 probe but not with AP2 oligo. The excess of

https://etd.uwc.ac.za/

the cold SP1 oligo blocked the formation of the protein-DNA complex (Fig. 3.5A). To identify whether the truncation Δ1-452 of the AcMNPV IE-1 binds to the AcMNPV *hr5*, the 60-bp direct repeat and the soluble fraction of the purified truncation Δ1-452 were employed to perform the gel retardation assay (Fig. 3.5B). When the truncation Δ1-452 was incubated with the ^{32}P labeled 60 bp probe, a single DNA-protein complex was formed and there was no non-specific complex detected. The addition of excess amount of the unlabeled 60 bp oligo with two different dosages inhibited the binding of the Δ1-452 to the ^{32}P labeled 60 bp probe, whereas the addition of an excess dose of the cold non-specific SP1 oligo did not block the formation of the protein-DNA complex (Fig. 3.5B).

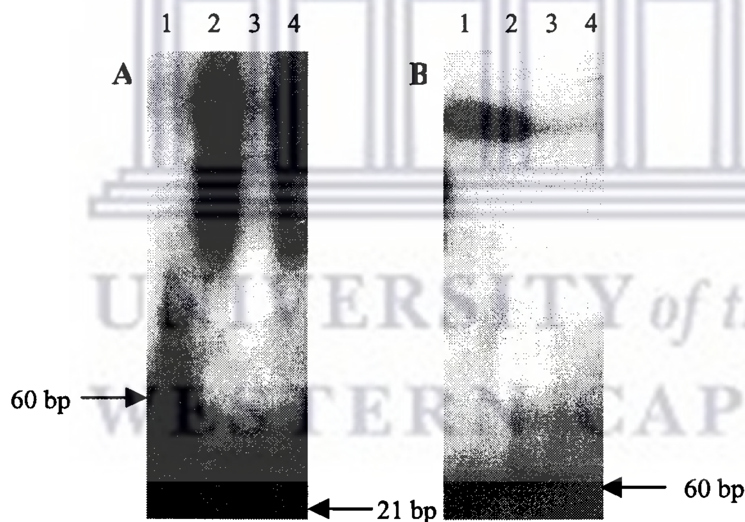


Figure 3.5 DNA-binding activities of the IE-1 truncation and controls. (A) EMSA of positive and negative controls. The 60 bp DNA probe did not interact with the fish viral His-tag fusion protein (lane 1). The SP1 oligo (21 bp) bound to a transcription factor in HeLa extract (lane 2). Prior to addition of the SP1 21 bp probe, excess amount of the

https://etd.uwc.ac.za/

competitor (cold SP1 21 bp oligo) and a non-specific AP2 oligo (26 bp) to the transcription factor were added to the reaction mixtures (lane 3 and 4). (B) Gel-retardation assay of the IE1 truncation. The IE1 Δ 1-452 interacted with the 60 bp probe in the absence and presence of the non-specific 21 bp SP1 oligo (lane 1 and 2). The non-radio labeled competitor (60 bp) with the different excess doses was incubated with the IE1 Δ 1-452 before addition of the 60 bp probe (lane 3 and 4).

3.5 Discussion

Oligomerization is believed to be required for many HLH-containing transcriptional activators to bind to palindromic DNA recognition sites (Murre *et al.*, 1989; Massari and Murre, 2000; Patikoglou1and Burley, 1997). This also occurs in baculovirus IE1 mediated transactivation (Olson *et al.*, 2001). Current evidence indicates that IE1 oligomerization occurs prior to *hr* binding (Olson *et al.*, 2001) and IE1 binds to *hr* as a dimer (Rodems and Friesen, 1995). Furthermore, the transactivation is directly correlated with the capacity of IE-1 to bind to the palindromic 28-mer, a *hr* enhancer element (Guarino and Dong, 1994; Kremer and Knebel, 1998; Rodems and Friesen, 1995). Therefore, the oligomerization of IE1 contributes to the DNA binding and transactivation function.

The previously functional mapping of the AcMNPV IE-1 demonstrated that deletion of 25 or more amino acid residues from the C-terminus abolished its DNA binding and enhancer-dependent transactivation function (Kovacs *et al.*, 1992). According to the

https://etd.uwc.ac.za/

inspection in this study, the 25 amino acid (aa) deletion actually removed almost half of the IE-1 basic helix-loop-helix (b-HLH)-like domain. On the other hand, the insertions at the different position of the b-HLH-like domain greatly reduced or eliminated the capacity of each mutated IE1 to bind DNA as a dimer (Olson *et al.*, 2001). For example, a series of insertions constructed at the positions of the eight hydrophobic residues (Fig. in a failure of 3.2B) within the HLH-like domain, caused the loss of both homo- and heterodimers, thus resulting transactivation, while the replacement of basic residues 537 and 538 in the domain appeared to exhibit a similar impact on oligomerization and transactivation *in vitro* to that which occurred in the cases of the insertion mutations of the hydrophobic residues. This event in which multiple residues were involved in oligomerization in such a small region was not observed in other N- and C-terminal regions. The findings revealed that the b-HLH-like domain of the IE1 indeed is critical for oligomerization and DNA binding. The comparative analysis of the basic HLH-like domains in the present study indicated that the domains are highly conserved in the putative primary and predicted secondary structures of the C-terminal half of those IE-1s (Fig. 3.2A). This suggested the event occurring in the b-HLH-like domain of the AcMNPV IE1 may possibly happen in the case of other NPV IE1s.

UNIVERSITY of the
WESTERN CAPE

As suggested previously, the multiple regions may be involved directly or indirectly in DNA-binding or oligomerization (Rodems and Friesen ,1995; Rodems *et al.*, 1997). To investigate whether a b-HLH-like domain-containing region could bind to DNA independently, a deletion mutant Δ 1-452 was used in the binding assay, (where only the b-HLH-like domain was identified to contribute to IE-1 oligomerization). Although it

https://etd.uwc.ac.za/

was reported that an insertion at the conserved residue 425 exhibited a disruption of DNA binding (Rodems *et al.*, 1997), due to exclusion of the residue from Δ 1-452, the binding function of the b-HLH-like domain could be distinguished from that of the residue or other regions of IE-1 which are possibly involved in DNA-binding and oligomerization. On the other hand, several hydrophobic residues which are required for oligomerization and contribute to IE-1 stability (Olson *et al.*, 2001), were centered in the HLH-like domain which has about 50 aa in size. Such a structure was not found in other binding- and oligomerization-related regions of the AcMNPV IE-1 where only certain specific residues were identified as being directly or indirectly involved in oligomerization and DNA binding (Olson *et al.*, 2001; Rodems, *et al.*, 1997). The binding activity of Δ 1-452 showed that the HLH-like domain-containing truncation could bind to the 60 bp repeat directly and specifically. Therefore, it is more likely that the b-HLH-like domain could be mainly responsible for the direct binding. However, except for Δ 1-452, whether or not other regions of the IE1 could also bind to the repeat independently, was not excluded. The analyses of both the binding result and the structural comparisons of the C-terminal regions of the IE1s suggested that the binding activity of the AcMNPV IE1 deletion mutant (Δ1-452) may remain in the corresponding region of other NPV IE1s, including TnSNPV IE-1. Furthermore, expression and detection of the TnSNPV IE-1 Δ 1-646 (93 aa) (Fig. 3.3) suggests that it is possible to further reduce the binding region down to a smaller size such as 93 aa. This study also provided the purified protein sources for identifying the HLH-like domain by NMR or X-ray crystallograghy studies.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

The previous binding experiment of IE1(Guarino and Dong, 1991) indicated that two protein-DNA binding complexes were formed when a ^{32}P-labeled 252 bp fragment consisting of the left half of *hr5* was used. However, when a single 60 bp repeat within the left half fragment was applied, only one complex was formed in the experiment. It was suggested that this was because IE1 protein bound simultaneously to one or two copies of a 60 bp direct repeat (DR60) contained in the *hr*5 fragment. In this study to eliminate the impact of possible cellular binding factors on the binding reaction by protein purification, and using a 60 bp probe only, the result was consistent with the previous one.

Current evidence showed that a 28 bp palindrome was a minimal sequence for DNA binding (Rodems and Friesen, 1995). It represented the smallest functional unit of *hr5* capable of position- and orientation-independent transcription activation (Rodems and Friesen, 1993). However, a functional dissection of the left half (252 bp) of the *hr*5 revealed that a conserved 24-bp palindrome within the 60 bp direct repeat (DR60) was mapped as a minimum sequence for DNA-protein interaction, but a complete copy of the DR60 was required for enhancer function (Guarino and Dong, 1994). The conclusion was different from that in which a 28 bp repeat was essential for DNA-protein interaction and enhancer-dependent transactivation of IE1 (Rodems and Friesen, 1995). In accordance with the inspection of the entire *hr5* sequence (Margot *et al*., 1992) in this study, only half of the 28 bp palindrome, referred to as *hr5*-A, was found in the 252 bp fragment (Guarino and Dong, 1991) and its sequence differed from that of the 60 bp repeat (*hr5*-B) in which the 24-bp palindrome is present. According to the differences of

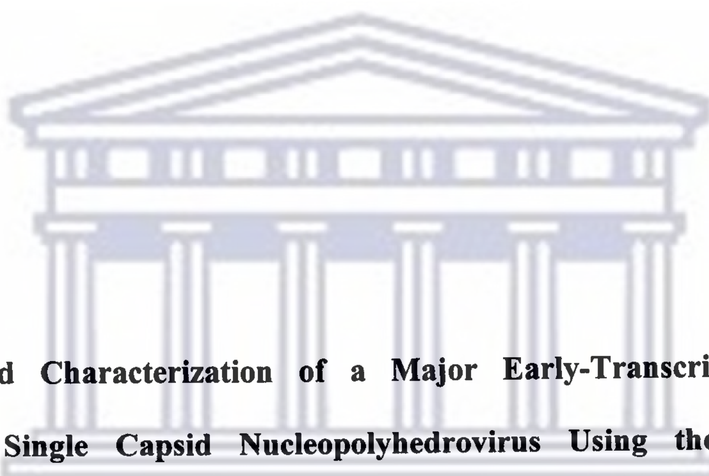
UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

the locations and the sequences of 28 bp (*hr5*-A) and DR60 (*hr5*-B) in *hr5*, it appears that there are two different essential repeats in the left half of the *hr5* which are capable of DNA-protein binding and enhancement of transactivation. It will be interesting to investigate whether there is any functional difference between the two independent enhancer elements.

In summary, an independent DNA-protein binding region of the AcMNPV IE-1 which contains a b-HLH-like domain was identified. Although this result did not define the smallest protein binding unit of the IE-1, it indicated that there is a direct binding domain in the 130 amino acid region. This study also suggested that the HLH-like domain in the binding region could be mainly responsible for the DNA binding as several residues involved in oligomerization were centered in the domain. It will be interesting to investigate the functional role of the domain further by analyzing the domain fusion protein. Due to the structural similarity of the C-terminal binding region in AcMNPV and other NPV IE-1s, the protein-DNA binding reaction may occur, not only in the case of AcMNPV IE-1, but in the corresponding regions of the TnSNPV and other NPV IE-1s. The study also provided the purified protein source for studying the structure of the b-HLH-like domain of IE-1 by NMR or crystallography.

https://etd.uwc.ac.za/

Chapter 4


UNIVERSITY of the
WESTERN CAPE

Identification and Characterization of a Major Early-Transcribed Gene of *Trichoplusia ni* Single Capsid Nucleopolyhedrovirus Using the Baculovirus Expression System

https://etd.uwc.ac.za/

4.1 Abstract

An early transcribed gene (*me-53*) of *Trichoplusia ni* single capsid nucleopolyhedrovirus (TnSNPV) was sequenced and identified. It has an open reading frame of 1146 nucleotides that encode a protein of 382 amino acids with a molecular mass of 45.2 KDa. Upon introduction of the gene and a green fluorescent protein (GFP) reporter gene into the baculovirus expression vector system (BEVS), the transcriptional analysis of the gene in two cell lines infected with the *Autographa californica* nucleopolyhedrovirus (AcMNPV) recombinant revealed that an early transcript can be detected at 1h postinfection (p.i.) until 12 h p.i., and a late one from 18 h p.i. up to 48 p.i. A unique early transcription start site was identified at two bases upstream of a conserved late transcription start site TAAG. The putative regulatory elements within the upstream regulatory regions of ten NPV *me-53s* were also identified and discussed. The protein analysis showed that zinc finger-like motifs at the C-termini of twelve baculovirus ME-53s are highly conserved with similar patterns. The phylogenetic analysis of these ME-53s indicated that the TnSNPV ME-53 is more closely related to the members of NPV subgroup II-A, *Spodoptera exigua* multiple (SeM) and *Mamestra configurata* (Mc) nucleopolyhedrovirus ME-53s. Due to two heterologous genes (TnSNPV *me-53* and GFP) introduced into the AcMNPV polyhedrin locus, the recombinant was applied to test a new model system for improvement of bio-insecticidal agents for insect pest control.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

4.2 Introduction

Baculovirus gene expression is believed to be regulated in a cascade fashion in which each successive phase precedes the previous phase (Blissard and Rohrmann, 1990). An early phase occurs prior to viral DNA replication and late phase during or after the viral DNA synthesis (Blissard and Rohrmann, 1990). The early transcription is mediated by host RNA polymerase II (α-amanitin-sensitive) (Fuchs *et al.*, 1983; Hoopes and Rohrmann, 1991) and independent of viral replication and late gene expression (Friesen, 1997). Many baculovirus early promoters contain a conserved sequence CAGT at or near the transcription start site (Theilmann and Stewart, 1991 and 1992; Krappa and Knebel, 1991; Blissard and Rohrmann, 1989) and a functional TATA box (Blissard *et al.*, 1992; Theilmann and Stewart, 1991; Dickson and Friesen, 1991; Guarino and Smith, 1992). The two *cis*-regulatory elements, TATA and CAGT are essential for start site selection and initiation efficiency of early genes respectively, and may functionally work in a cooperative manner (Blissard *et al*, 1992). On the other hand, some sequence elements in the baculovirus *gp64* may be involved in transcription initiation by a TATA-independent mechanism (Kogan *et al*, 1995). Except for the early transcription, some of the early genes, with a conserved late transcription motif (A/G/T)TAAG (Friesen, 1997), such as the *me-53* and *p35* (Knebel *et al.*, 1993; Nissen and Friesen, 1989), can be recognized and transcribed by a α-amanitin-resistant virus-encoded RNA polymerase (Guarino *et al.*, 1998) in the late phase of viral infection, and others, such as the *ie-2, pe38* and *he65*, have the same conserved motif which has no function display during the late phase of infection (Carson *et al.*, 1991; Krappa and Knebel, 1991; Becker and Knebel, 1993). The

https://etd.uwc.ac.za/

evidence from a study indicated that the differential selection of TAAG sites was determined primarily by the immediate context of flanking sequence, rather than the position responsible for TAAG utilization (Garrity *et al.*, 1997).

To date 5 very early genes of baculoviruses have been identified. They are *ie-0, ie-1, ie-2, pe-38* and *me-53* genes (Kovacs *et al.*, 1991; Theilmann and Stewart, 1991; Carson *et al.*, 1988; Krappa and Knebel, 1991; Knebel *et al.*, 1993). Among them the first four play a regulatory role in either transactivating other genes (Nissen and Friesen, 1989; Kovacs *et al.*, 1991; Guarino and Summers, 1986c; Chisholm and Henner, 1988; Carson *et al.*, 1988 and 1991; Lu and Carsten, 1993) or regulating its own promoters (Carson *et al.*, 1991) as well as being involved in negative regulation of transcription (Kovacs *et al.*, 1991; Leisy *et al.*, 1997), while the AcMNPV *me-53* gene has been only postulated to play a regulatory role in the course of infection. This has been based on a suggested DNA binding domain in the deduced amino acid sequence of the gene (Knebel *et al.*, 1993). However, there is a lack of experimental evidence to identify the regulatory role of *me-53* in baculovirus gene expression and therefore, its functional role in gene regulation still remains under investigation.

UNIVERSITY of the WESTERN CAPE

The available lepidopteran baculovirus genomic sequence data from GenBank has shown that *me-53* homologs are present in all these viral genomes. However, until now only the AcMNPV *me-53* gene was characterized and the rest of me-53s in different baculoviruses has not been experimentally identified. To identify the TnSNPV *me-53* gene, a green fluorescent protein (GFP) reporter gene and the TnSNPV *me-53* under the control of its

https://etd.uwc.ac.za/

own promoter were introduced into the polyhedrin (*polh*) locus of a commercially available AcMNPV shuttle vector (bacmid) by site-specific transposition, the system of which is referred to as a baculovirus expression vector system (BEVS) (Luckow *et al.*, 1993). The other purpose of using BEVS in this study was to test a model system of early expression of multiple foreign genes for improving the bio-insecticidal agent. On the other hand, there are some advantages for using the baculovirus expression system in this study: 1) It is relatively easy to generate a virus recombinant that carries a reporter gene, thus allowing the monitoring of transfection of recombinant bacmid and virus infection, as well as easily titering the occlusion body minus (occ^-) recombinant; 2) It permits the rapid and simultaneous isolation of multiple recombinant viruses without the time-consuming multiple rounds of plaque purification; 3) Due to the foreign gene entering cells through viral infection the gene copies of the expression will be much greater than those by plasmid transfection. This is especially suitable for early genes without a strong promoter.

UNIVERSITY of the
WESTERN CAPE

Previously, polyhydrin, *p10* and *ie-1* homologs of the South Africa TnSNPV isolate, which is highly pathogenic to *Trichoplusia ni* larvae, were identified (Fielding and Davison, 1999; Fielding and Davison, 2000; Wang *et al.*, 2001; Fielding *et al.*, 2002). In these studies a few attempts were made to contribute to the placement of the virus in baculovirus phylogeny. Although the estimate of the virus genome size is about 160 kb, proximately 30 kb larger than that of AcMNPV (Fielding *et al.*, 2002), there is a lack of information of the entire virus genome. Thus, it is necessary for phylogenetic analysis of the virus to compare combining data from multiple genes to reduce sampling errors and

https://etd.uwc.ac.za/

allow phylogenies to converge towards correct solutions with good support (Mitter *et al.*, 2000). A *me-53* gene, like an *ie-1,* is well conserved among the lepidoptaren baculovirus genomes. As a member of the early gene class, the TnSNPV *me-53* can be considered as another candidate, together with the ie-1, for inferring TnSNPV phylogeny because of its relatively conserved and longer amino acid sequence in comparison with the previously identified polyhedrin gene that has been thought not to be the ideal gene for phylogenetic analysis, due to its small size and high sequence conservation (Bulach *et al.*, 1999).

The present study demonstrated that the TnSNPV *me-53*, as a heterologous gene, was faithfully transcribed as the authentic *me-53* did in the context of the AcMNPV recombinant. The early transcription initiation of the gene occurred from a unique mRNA start site at a very early phase of the viral infection and the late promoter of the gene can also be recognized by the virus-encoded RNA polymerase at the late stage of viral replication. In addition, the upstream regulatory region analysis of ten NPV *me-53*s was conducted and the deduced amino acid sequence of the TnSNPV ME-53 was also compared with other baculovirus ME-53s to identify common features.

UNIVERSITY of the
WESTERN CAPE

4.3 Materials and Methods

4.3.1 DNA manipulation of the TnSNPV *me-53*, sequencing and computer analysis

Appropriate templates containing the TnSNPV *me-53* in an 11 kb fragment from a TnSNPV *EcoR*1 genomic library were prepared by exonuclease III digestion (Henikoff 1984). Sequencing was conducted using the Sequitherm kit (Epicentre Technologies)

https://etd.uwc.ac.za/

with CY-5 labeled primers. Nucleotide sequence was resolved on an Alfexpress automated DNA sequencer (Pharmacia). The final sequence was confirmed in both the sense and anti-sense directions (Fig. 4.1A).

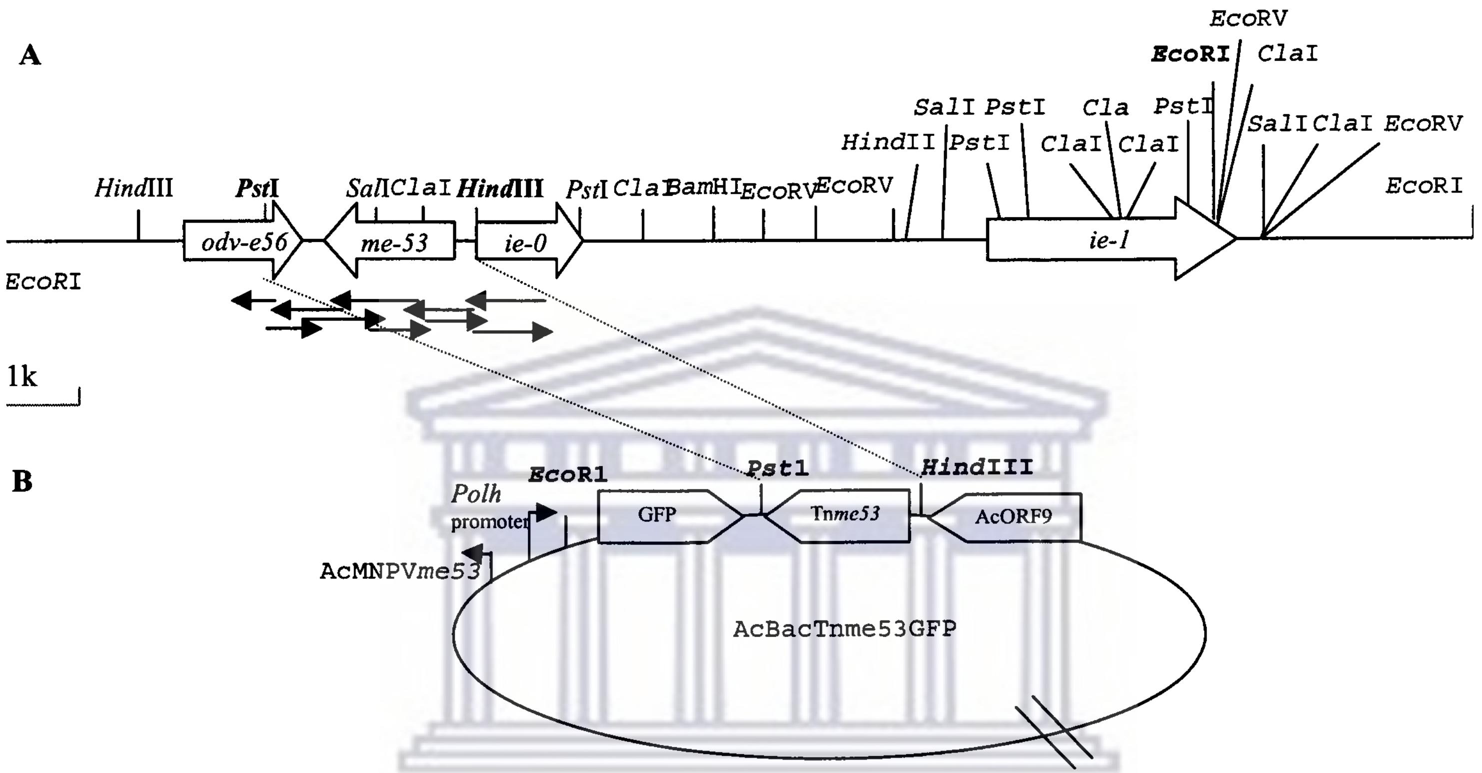


Figure 4.1 The location of the *me-53* gene within an 11 kb and an adjoining 2.3 kb *Eco*RI fragment and the AcMNPV recombinant (Bacmids). (A) The *Pst*I and the *Hind*III sites used for the *me-53* subcloning within the 11 kb fragment are in bold. The sequencing directions are represented by arrows. (B) The orientation and location of relevant genes in the recombinant virus are indicated by arrows. The restriction sites for cloning TnSNPV *me-53* and GFP into pFast Bac HT expression vector for site-specific transposition are in bold. The relation of the *me-53* gene within the11kb fragment to that in the recombinant (AcBacTnme53GFP) is shown by dashed lines. The structure of the control recombinant

https://etd.uwc.ac.za/

AcMNPV/GFP (AcBacGFP) is identical to that of AcBacTnme53GFP, excluding the TnSNPV *me-53* (diagram not shown).

Nucleotide and amino acid sequence manipulation was carried out using the University of Wisconsin, Genetics Computer Group GCG sequence analysis package. Multiple sequence alignments were conducted using the ClustalW program (Thompson *et al.*, 1994). Based on alignment data sets of the baculovirus ME-53s and IE-1s, phylogenetic trees were constructed, using the neighbor-joining method implemented in the Clustal W program, and the robustness of the data sets was tested by bootstrap analysis conducted in TreeView package (Win32). Protein comparisons with entries in the updated GenBank/EMBL and SWISS-PROT databases were performed with BLAST program (Altschul *et al.*, 1997). Alignment editing was conducted with Genedoc Software. ME-53s were analyzed using the ExPASY server (Appel *et al.*, 1994) for prediction of domains and motifs. The sequences of the 5' non-coding regions of the baculovirus *me-53*s were obtained from GenBank database, except for those of the TnSNPV and CfMNPV *me-53*s. The sequence of the CfMNPV *me-53* gene was kindly provided by Dr. Krell (University of Guelph, Canada). Relevant regulatory elements in the regulatory regions of *me-53s* were located and the same elements were manually grouped.

4.3.2 Cells and virus

S.frugiperda 9 and High5 cells were maintained at 27/ C in Grace medium supplemented with 10% fetal calf serum (Gibco). Three passage stocks of the recombinant viruses AcMNPV bacmid /TnSNPV *me-53*/GFP (AcBacTnme53GFP) and AcMNPV

https://etd.uwc.ac.za/

bacmid/GFP (AcBacGFP) were inoculated at a multiplicity of infection (MOI) of 10 p.f.u. per cell.

4.3.3 Generation of the recombinant baculoviruses

A H*ind*III-*Pst*I fragmant containing the TnSNPV *me-53* in a 11 kb fragment (Fig. 4.1A) from a TnSNPV *EcoR*1 genomic library was cloned into the PUC18 vector (i.e. PUC18/*me-53* plasmid). The pEGFP-N1 vector (GenBank accession number: U55762) carrying a green fluorescent protein gene (GFP) under the control of Human cytomegalovirus (CMV) immediate early promoter (Clontech) was digested with *Hind*III and *Bam*H1 to remove *Eco*R1 and *Pst*1restriction sites from its multiple cloning site and then the sticky termini were filled with klenow and ligated by T4 DNA ligase (Invitrogen). Two primers introduced with *Eco*R1 and *Pst*1 sites (pEGFPN1F: 5'-ctctctctgaattc-TAGTTATTAATAGTAATCAATTACGGGGTC-3' and pEGFPN1R: 5'-tctctctcctgcag-CATTGATGAGTTTGGACAAACC-3') were designed to amplify CMV promoter and GFP from the modified vector using *pfu* DNA polymerase. The resulting PCR products were digested with *Eco*R1 and *Pst*1 and then inserted into the *Pst*1-*Eco*R1 sites of the PUC18/*me-53* plasmid in the opposite orientation to the TnSNPV *me-53* (i.e. PUC18/*me-53*/GFP) (Fig. 4.1B). The *Hind*III-*Eco*R1 fragment containing *me-53*/GFP digested from the PUC18/*me-53*/GFP plasmid and the previous purified PCR products (containing the GPF only) with the *Eco*R1 and *Pst*1 ends (for constructing the control recombinant bacmid) were inserted into the $pF_{AST}B_{AC}$ HTb donor plasmids, respectively. The donor plasmids were transformed into $DH10B_{AC}$ cells, white colonies carrying the recombinant bacmids were selected, cultured and finally the bacmids were purified

https://etd.uwc.ac.za/

according to manufacturer's protocol (Qiagen). The presence of the AcMNPV and the TnSNPV *me-53*s in the bacmids were confirmed by PCR (date not shown). The AcMNPV/TnSNPV*me-53*/GFP and AcMNPV/GFP recombinants, AcBacTnme53GFP (Fig. 4.1B) and AcBacGFP (control), were generated by transfecting Sf9 cells with the two corresponding purified bacmids, respectively, using TransIT-Insecta Reagent as recommended by the manufacturer (Mirus). The virus stocks were amplified for two more rounds, following Qiagen's protocol. The transfection of the viral bacmids, and each of the viral infections, were confirmed by monitoring expression of GFP under a fluorescence microscope after harvesting viruses at 72 and 48 h p.i., respectively. The titer of each of the occ$^-$ viral stocks was determined by the end-point dilution method (O'Reilly *et al.*, 1992) under a fluorescence microscope.

4.3.4 Total RNA isolation, 5' and 3' RACE and RT-PCR analyses of the TnSNPV *me-53* transcripts

Total RNA was isolated from 2 x 10^6 mock-infected and AcBacTnme53GFP-infected High5 cells at 1, 3, 6, 9, 12, 18, 24, 36 and 48 h p.i. The total RNA was also isolated from the same density of the AcBacTnme53GFP-infected Sf9 cells at 3 and 24 h p.i. and the AcBacGFP-infected High5 cells at 48 h p.i. TRIzol reagent was used for the RNA isolation as described in the manufacturer's protocol (GibcoBRL).

The primers used for 5' end mapping of the authentic *me-53* of AcBacTnme53GFP (Ac384R: 5'-CGTGGCTTTCATGATGTCG-3'; Ac317R: 5'-ATCCAGTTTGTCGCTGT-ACG-3'; Ac230R: 5'-ATGCTGCTGTGCTCCAACG-3') were designed from the

https://etd.uwc.ac.za/

previously published AcMNPV *me-53* sequence data (Genbank accession number: L05914). The primer used for 5' and 3' end mappings of the TnSNPV *me-53* were shown in Fig. 4.2. The total RNA isolated from AcBacTnme53GFP-infected High5 and Sf 9 cells was used to perform the 5' end mappings of the authentic *me-53* transcripts (at 3 and 24 h p.i. for both cell lines) and the TnSNPV *me-53* transcripts (from 1 h to 48 h p.i. for High5 cells and 3 and 24 h p.i. for Sf9 cells), and the 3' end mapping of the TnSNPV *me-53* transcript (at 3 and 24 h p.i. for High5 cells). The 5' and 3' RACE for the mappings were carried out as described by Sambrook and Russell. (2001).

```
1     TTGACGAATAGAATGATAGTTTTTCTTATTTGTTTTAGATACGGCATAAATACGTACATA
          V  F  L  I  I  T  K  R  I  Q  K  L  Y  P  M  F  V  Y  M

61    CTATATTCAAACTTTGACATGCGATTAACGTGCTGTAAAACATTAACGATTTTACCGATT
       S  Y  E  F  K  S  M  R  N  V  H  Q  L  V  N  V  I  K  G  I

121   TCCGCTATCAAATAATGTAAACATCTATCGTTGGGAATACGAATGTCGTCAGTGTCATCT
       E  A  I  L  Y  H  L  C  R  D  N  P  I  R  I  D  D  T  D  D

181   ACAAACGCCAGAATCTTGTCGATTTTTTCGTTATATGTCTGTTGAAAGATCTCGTCTATC
       V  F  A  L  I  K  D  I  K  E  N  Y  T  Q  Q  F  I  E  D  I

241   ACATTAAATGAGGCCTGTTTAACTAAAAATTGGGCCTTGGGATTCAACGCTAAATCGTTG
       V  N  F  S  A  Q  K  V  L  F  Q  A  K  P  N  L  A  L  D  N

301   GTGTGCAGGTCGTTAAAAATAAAATTTTTAAATACCTGTGTTTCGAAATCGGTCGAGCCT
       T  H  L  D  N  F  I  F  N  K  F  V  Q  T  E  F  D  T  S  G

361   TTTTCTCCATTGTTGTTTCTAATGTTCAAGAAACTATTTATACTATTACTGTTTCCATCG
       K  E  G  N  N  N  R  I  N  L  F  S  N  I  S  N  S  N  G  D

421   TCAGAGTTTCGCGGTGACAATAGAATGCTGGCCACAGCCGCACTTTCCATTGCGATACGC
       D  S  N  R  P  S  L  L  I  S  A  V  A  A  S  E  M  A  I  R

481   TTATAAAGATCGTTTTGATAATCAAAAGTGTTCATGACAATCGCCCGATGATGACTGACC
       K  Y  L  D  N  Q  Y  D  F  T  N  M  V  I  A  R  H  H  S  V
                        HindIII
541   ACAACATCGTATGAGATAAGCTTTTGTACTGGGTCCATGACTACAAAATCACGTACCAGA
       V  V  D  Y  S  I  L  K  Q  V  P  D  M
                                    <---  Ie-0
                                                            -162
601   CAGGATGAGCATCTTCATGTAAATGAATCGCGCTATTGTCTCGCGTCTCGTAATAATATA

                                                            -101
661   AACTTGACGGGTGACGCCTTTGTAGAGTCAAATTCCCGCGCATATTATGTGTTTATTTAT
```


https://etd.uwc.ac.za/

```
                                             -64
721   AACTCTACGGTCGCGCCGAGGTCAACAGAGCGTGTATAAAGTGTTCGCTTCTCGTATTTA
        -34  -31
             Me-53                           +1       Tn241F
781   TTTTCATAAGCTATCGTGCGCTTCTAAACATCGGTCATATGTCATCGACGACCCAATCGT
                                              M  S  S  T  T  Q  S  Y

841   ATTCCACGACAACAGTCAAAATGTCTTCAAAAACGGAAAGGAAACCCATGCTGATAGACC
       S  T  T  T  V  K  M  S  S  K  T  E  R  K  P  M  L  I  D  H

901   ACCTGATGAATATAGTTCGAGAAACTAGAGCCAACAATCCAGATTATAGACAGCCGCCCA
       L  M  N  I  V  R  E  T  R  A  N  N  P  D  Y  R  Q  P  P  I

961   TCATTCAAAGCCCTCCGAAGGATTTGAGACCTAAATTCTTGAGTCATTTGAATTATTGTC
       I  Q  S  P  P  K  D  L  R  P  K  F  L  S  H  L  N  Y  C  L

                                       Tn241R
1021  TACTACAATCGATAATTCAGTTTGCGTCCGATTACGTTCAGGGAAGATTAAAAGTGAACA
       L  Q  S  I  I  Q  F  A  S  D  Y  V  Q  G  R  L  K  V  N  N
                          Tn293R
1081  ATTTGCTACAGATGAATTGTAGAGATTTGAAGCAAGAAGAAATTGTAGAAAAGGCTACAT
       L  L  Q  M  N  C  R  D  L  K  Q  E  E  I  V  E  K  A  T  C
                                                                ▲
       Tn333R
1141  GTAACCAATGCTCGAAACAGTTTAAGCGAGAGACTAATTTCTCACTCTTTTGCCTCATCG
       N  Q  C  S  K  Q  F  K  R  E  T  N  F  S  L  F  C  L  I  D
             ▲

1201  ACAACAGCATAGACATTGAAGATGAATTAAAGAGACACAAGAAGTTTAAATTGATCTGCT
       N  S  I  D  I  E  D  E  L  K  R  H  K  K  F  K  L  I  C  S
                                                             ▲

1261  CTAAATGTTACGAACTGTTTAAATATAAACACCAGTTTGACCTGTTTCAGATATATCCAT
       K  C  Y  E  L  F  K  Y  K  H  Q  F  D  L  F  Q  I  Y  P  Y
          ▲

1321  ATGTTATACTCGAAGAAGCAGAGACTCTGTGTCAATTGGGCTTTTTTAAATGCTACTTGT
       V  I  L  E  E  A  E  T  L  C  Q  L  G  F  F  K  C  Y  L  F

1381  TCAATATTAATCTCGAACACACTTGTACTACAGAAGAGATTTCGGTTGTGGGTCGACACG
       N  I  N  L  E  H  T  C  T  T  E  E  I  S  V  V  G  R  H  D

1441  ATTTCTTTGGGACAATAAAGGATATAGTGGATAAAAAAAAGCCCAACGAACAAATCACAA
       F  F  G  T  I  K  D  I  V  D  K  K  K  P  N  E  Q  I  T  K

1501  AAATTTTATTGCGAACATATGGTCGAACTCTATTTACAGAAACTGACAGAAATTGCATCA
       I  L  L  R  T  Y  G  R  T  L  F  T  E  T  D  R  N  C  I  I

1561  TAAAGACTACTTCTAGCAAGATCGGCTGCGATGACAACACTTTTCAGCTATATTTCGGCG
       K  T  T  S  S  K  I  G  C  D  D  N  T  F  Q  L  Y  F  G  D
                                                      Tn313F      Tn297F
1621  ATAGCAAAATGATGGATTTTTTTGCGAGCATGGGCGAACAGAAGCTGCTCACCTATTTCT
       S  K  M  M  D  F  F  A  S  M  G  E  Q  K  L  L  T  Y  F  Y

1681  ATTCGGTCGAAAAAAAAAGTCTACAAAACTACTTTTAATTTTGTACTATATTTTCCTATAC
       S  V  E  K  K  V  Y  K  T  T  F  N  F  V  L  Y  F  P  I  P

1741  CATGTAAGCGATTTTGCATTCTCTGCACTCGACATAAAATGTATCTCAAAAAACATATAG
       C  K  R  F  C  I  L  C  T  R  H  K  M  Y  L  K  K  H  I  V
       ▲           ▲     ▲

1801  TGCTGTATTGTTCTCAGTGCGGTTTTACGGATGCCATGTTTTTCACTAGAAACAAGCTCG
```

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

```
        L  Y  C  S  Q  C  G  F  T  D  A  M  F  F  T  R  N  K  L  D
              ▲        ▲
1861  ACGTGTCTGCTTTAAAATTTTATAGCGTTTGTGTTAAAGTTAAAACCATCAAGCCCAAAC
        V  S  A  L  K  F  Y  S  V  C  V  K  V  K  T  I  K  P  K  R

1921  GTATCTACTATTACGATATGAATCTTTATAAAAAAATTAATGTATAAAAATTATGTGTAT
        I  Y  Y  Y  D  M  N  L  Y  K  K  I  N  V  *

1981  AAATTATGTATATTCATTCGAAAGTCTAATAAAAATCCTTCGGGCTACACATAATTGTTT

2041  TATTATGAATCGCCATGACACTAAAACAGTCGTAAATCGTCTAGACGACAAAGTGACATG

2101  ATTAAGCCATTCAACGAGTTCTATTAGTTTCTACGGCTCACAGTTTGCTTATTGAACAGT
                                    *  N  R  R  S  V  T  Q  K  N  F  L

2161  CTTTTGATAACAAAATATCCAATGATAATTATAAATGCAATAACTCCTATCATTACAATA
        R  K  I  V  F  Y  G  I  I  I  I  F  A  I  V  G  I  M  V  I

2221  AACGGCAGTAAACTATCGCTTAAACTTTGACTTTTATTTGAAGATTTATTAATCAAGCCG
        F  P  L  L  S  D  S  L  S  Q  S  K  N  S  S  K  N  I  L  G

2281  TTTTCGCCCAATAAATTATCTAAATTCAAATCTCCTATTAAATCACCCATATTGTAGGGT
        N  E  G  L  L  N  D  L  N  L  D  G  I  L  D  G  M  N  Y  P

2341  TCTACACAGGCGATAGTTTGACCGGGAGCTAAGTCGGATATATCTACATATTGAGGTGTA
        E  V  C  A  I  T  Q  G  P  A  L  D  S  I  D  V  Y  Q  P  T

2401  TTTATGTTGGCGTTGGGGTCGCTAGCCCGACATACGGATTTTTCTAGTTCATAATTAAAG
        N  I  N  A  N  P  D  S  A  R  C  V  S  K  E  L  E  Y  N  F
                          PstI
2461  CCTTGACAAATACTCTGCAG
        G  Q  C  I  S  Q  L
                 ←—— Odv-e56
```

Figure 4.2 Nucleotide and predicted amino acid sequences of the TnSNPV *me-53*-containing fragment. The early and late initiation sites of the *me-53* are depicted by arrows. The TATA box-, CG- and CGT-like motifs and a CGTGC sequence in the upstream regulatory region and a polyadenylation signal downstream of the stop codon are underlined. The position of the transcription sites and TATA-like motifs, relative to the translation start codon, is indicated by numbers. The poly (A) chain attachment site is double underlined. The cysteine residues of a putative zinc finger motif at the C-terminus of the ME-53, and two pairs of the function-unknown cysteine clusters at the N-terminal half, are marked by solid triangles. Two histidine residues within the zinc finger-like

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

motif are highlighted. The gene-specific primers used for 5' and 3' RACE are underlined and the bases shared between two primers (Tn313F and Tn297F) are in italics. The transcription direction of three genes are shown by arrows. A putative polyadenylation signal of *odv-e56* gene is underlined. The restriction sites for cloning the fragment into PUC18 are in bold.

Briefly, for 5' RACE 14 μg total RNA per time point were used as templates to carry out reverse transcription (RT) using SUPERScript II reverse transcriptase (Invitrogen) with the gene specific primers Ac384R or Tn333R after DNase treatment and phenol clean-up. The first strand cDNAs were purified with the PCR purification kit (Qiagen) and a poly(A) tail was added to the 3'-ends of the cDNA using terminal transferase (Invitrogen) with dATP. Subsequent first and second PCR were performed to amplify the cDNAs using the oligo $(dT)_{18}$-anchor primer and the nested gene specific primers (Ac317R and Ac203R for the authentic *me-53*, and Tn293R and Tn241R for the TnSNPV *me-53*), respectively. The obtained PCR products were gel purified, cloned into pGEM-T vector and sequenced with M13 forward and reverse primers.

The total RNA of 5 μl isolated previously from AcBacTnme53GFP-infected High5 cells at 3 h and 24 h p.i., and the same oligo $(dT)_{18}$-anchor primer used in the 5' RACE assays were employed for 3' RACE as recommended by Sambrook and Russell (2001). The first strand cDNAs were amplified by PCR using the $(dT)_{18}$-anchor primer and the gene specific primer Tn313F. A second PCR was performed using the anchor primer and the nested gene specific primer Tn297F. The subsequent treatment of the resulting PCR products was the same as that described for 5' RACE.

https://etd.uwc.ac.za/

RT-PCR was performed by employing the first-strand of cDNA synthesized from the RT reaction of the TnSNPV *me-53* at each time point. The cDNA from each previous reaction was diluted to 10^{-2} and 1μl from each dilution was used in each PCR reaction with 23 cycles of amplification using the gene specific primers Tn241F and Tn241R. The total RNA isolated from the mock-infected and the AcBacGFP-infected High5 cells at 12 h p.i. was used for control RT-PCR with the same primers Tn241F and Tn241R. The final PCR products were resolved in a 1.2% agarose gel.

4.3.5 Phase contrast and fluorescence microscopy

Sf9 (1 x 10^5) and High5 cells (1 x 10^5) were infected by AcBacTnme53GFP with MOI of p.f.u. 10 per cell. The uninfected and infected cells at 48 h p.i were examined by phase contrast and fluorescence microscopy. The pictures of the bright field for uninfected and AcBacTnme53GFP-infected High5 and Sf9 cells were taken in phase contrast using Zeiss Axioscope with a 20x objective and a Sony DSC-S70 digital camera. Fluorescent pictures for the above infected cells were also taken using the same camera with a Zeiss Optivar filter set with an excitation wavelength of 450-490 nm and a 20x objective.

4.4 Results

4.4.1 Sequence analysis of the TnSNPV *me-53*

Analysis of the nucleotide sequence of the TnSNPV *me-53* revealed an open reading frame (ORF) of 1146 nucleotides, encoding a putative protein of 382 amino acids, with a

https://etd.uwc.ac.za/

molecular mass of 45.2 Kda (Fig. 4.2). The zinc finger-like motifs at the C-terminus are conserved among these ME-53s (Fig. 4.3)

```
AcMNPV  : MNRFFRENNIFDAPRTGGKGRAKSLP-APAANSPPSPVRPPPKSNIK :  46
BmNPV   : MNRFFRENNIFDAPKTGGKGRVKSLP-TPVANSPLSPVRQPPKSNIK :  46
CfMNPV  : MK-WFRENNIFDK-RSSRTAAADKP---AVATQGPAS---PAARRVK :  39
OpMNPV  : MN-WFKENNIFDN-KLAKKPAASKLGASPTASRPAGS---PAPRRVK :  42
McNPV   : ----------------------------------------------- :   -
SeMNPV  : ----------------------------------------------- :   -
HzSNPV  : ----------------------------------------------- :   -
HaSNPV  : ----------------------------------------------- :   -
TnSNPV  : ----------------------------------------------- :   -
LdMNPV  : ----------------------------------------------- :   -
SpltNPV : ----------------------------------------------- :   -
XcGV    : ----------------------------------------------- :   -

AcMNPV  : PPTRISPPKQPTRTSPAKPLEH---SSIVSKKPVVNRKDGYFVPPEF :  90
BmNPV   : PPTRIS---LPTRTFSANPLERSISSSIVSKKPVVNRKDGYFVPPEF :  90
CfMNPV  : PLNKSE-----Q--AHAAIIKR-----IGRGSDKLNDISASLVPPEY :  74
OpMNPV  : PLNKSE-----L--AHAAIVKR-----IGRGSDRLNEISASFVPPEY :  77
McNPV   : -------------------------------------------MRTV :   4
SeMNPV  : ---------------------------------MMKKTTTTTVISSS :  14
HzSNPV  : ----------------------------------------------- :   -
HaSNPV  : ----------------------------------------------- :   -
TnSNPV  : --------------MSSTTQSYSTTTVKMSSKTERKPMLIDHLMNIV :  33
LdMNPV  : ----------------------------------------------- :   -
SpltNPV : ----------------------------------------------- :   -
XcGV    : ------------------------------------------MTKRR :   5

AcMNPV  : GNKFEGLPAYSDKLDFKQE-RDLPMHFMSDLERDIMKATLKFSTNYI : 136
BmNPV   : GNKLESLPAYSDKLDFKQE-RDLRMHFMSDLERNIMKATLKFSTNYI : 136
CfMNPV  : GFRFDNVPACSHKLEYACE-RDLREHFLSDNEREAMKSLLRFATNYV : 120
OpMNPV  : GFRFDEVPACSHKLEYACE-RDLREHFLSDNEREAMKSLLRFATNYV : 123
McNPV   : KKNVTAKINTNIVPEVP---RDLRRKFLSVQNYNLLVSVSRFAADYA :  48
SeMNPV  : KKKVSAKVNTNRDDDDSKF-KDLPGRFLSNQNFQLMKYVLRFATNYV :  60
HzSNPV  : ----MASTSTAASLVNQHR-QDLRHKFLSVESKNLLCGMAKFADEYV :  42
HaSNPV  : ----MATTSTAASLVNQHR-QDLRHKFLSVESKNLLCGMAKFADEYV :  42
TnSNPV  : RETRANNPDYRQPPIIQSPPKDLRPKFLSHLNYCLLQSIIQFASDYV :  80
LdMNPV  : ----------MMSADAPP--RDSRREWLSEENAYLMSAVWRFAKDYF :  35
SpltNPV : ---------------MAGRIPDCRIRFMSFENRNLLEFAYNYASNYT :  32
XcGV    : SVVKKCEYITTIYAGTTRQTMDKRDQFLSQELRHALMFVVDFAKNAV :  52

AcMNPV  : MCYINSKDMRMTGKFASRPVKYKKT-MEHMSDSRCT-TCNYRFKDN- : 180
BmNPV   : MCYINSKDMRMTGKFASSSVKYKKTTEQHMSDSRCT-TCNYRFKDN- : 181
CfMNPV  : LCYVNSKDMLTFGRAAN--LKTKDG-LEHVQESECT-MCGYKFKEN- : 162
OpMNPV  : LCYINSKDMLTFGRAAG--LKLKNE-LEYVQESECT-MCGYKFKDN- : 165
McNPV   : RCRLRLNNLKLLGYDNN------NR-VENTKMTKCQGSCDLFFGPH- :  87
SeMNPV  : QCTYLVNDLKEMECTDL------KY-GEVERRTVCQGGCGKKFTGG- :  99
HzSNPV  : RCIHNVTQVNLHNCENL------KSPHDLAVRTMCD-KCQTVFRGPP :  82
HaSNPV  : RCIHNVTQVNLHNCENL------KSPHDLAVRTMCD-KCQIVFRGPP :  82
TnSNPV  : QCRLKVNNLLQMNCRDL------KQ-EEIVEKATCN-QCSKQFKRET : 119
LdMNPV  : LCVYRINDLFAMNCHQL------KNHHDEVARTTCD-RCKRRFGDPA :  75
SpltNPV : KCANDEN----------------------------CD-MCHRFFHDLP :  51
XcGV    : SCHSNVLRFTP-----------------------TCH-KCGKTFNQIY :  76
                                            △    △
```


https://etd.uwc.ac.za/

```
AcMNPV  : ----TREWFLYVVV-H-------IEKP-----LDDPDRIDICCQKCY : 210
BmNPV   : ----TRAWFLYVVV-H-------IEKP-----LDDSDRIDICCQKCY : 211
CfMNPV  : ----TRVWMLYVIVRHPPRSLSASEEPAAPPSPDAPGHFEFACCECA : 205
OpMNPV  : ----TRVWMLYVIVRHPPRAASADEE--FVPSPNTPDCFEFACCDCA : 206
McNPV   : ----CKVIYCVIDL--------EVDLD--DSATRWN-KFLMMCNNCA : 119
SeMNPV  : ----NKLLFCVIDN--------NGDDD--DVDDGGNEKKKVYCHVCS : 132
HzSNPV  : -F--TRWLFCAVNFR--------ISFD--NTKQKRDQKFKLVCEDCA : 116
HaSNPV  : -F--TRWLFCAVNFR--------ISFD--NTKQKRDQKFKLVCEDCA : 116
TnSNPV  : ----NFSLFCLIDN--------SIDIE---DELKRHKKFKLICSKCY : 151
LdMNPV  : HA--LEGLYCLANN--------KIVHE---SESNFHNRFKLICKPCC : 109
SpltNPV : -----RWLYCLVDK-----------------TN-TERFLFVCVDCW :  74
XcGV    : VDTHNSYMFMVVKS-----------WL-----QMDVDKVKFSCLQCK : 107
                                                      △  △

AcMNPV  : LYHN----VPKTSYEIYPSINIVDLSYLARERFFYQYIFPVS----L : 249
BmNPV   : LYHN----VPKTSYEIYPSINIVDLSYLARERFFYQYIFPVS----L : 250
CfMNPV  : DNYH----DQLNSHQVYPGISSVHAQRLFKSCFFYQYVFPLEF--RL : 246
OpMNPV  : DNYP----DQLNSHQVYPGINSLHAQRLVEACFFYQYVFPLEY--KT : 247
McNPV   : SDYKGD--DRYNVLQLFPHLKIVNVERLCALCFLTKYIFPIN----L : 160
SeMNPV  : NDYAYK--RRYEVMQLFPTVSISVVERLCEVCFITKYIFPID----L : 173
HzSNPV  : QTYILH--PEFQVYELYPRIHLKHVLELCRHCFIRKYFLPINPDLYS : 161
HaSNPV  : QTYILH--PEFQVYELYPRIHLKHVLELCRHCFIRKYFLPINPDLYS : 161
TnSNPV  : ELFKYK--HQFDLFQIYPYVIIEEAETLCQLCFKCYLFNIN----L : 192
LdMNPV  : AKVVDV--PTVELRQLYPRLDFDTVEWLARCRFVTRYIFPVE----T : 150
SpltNPV : KISVRDESKRYKNFELYPRVCLSDVWTLCKFGVFKKFHFKID----- : 116
XcGV    : NDPM------EDVIELYPSFSLANLKKLMYSGTLKRFVFPFI----- : 143

AcMNPV  : EHTTEVKELRIDDHNCKVFEIIRRIIRNHKEPNERIQTIDLSTTGGL : 296
BmNPV   : EHTTEVKELKIDDHNCKVFEIIRRIIRNHKEPNERIQTIDLSTTGGL : 297
CfMNPV  : EHFT-FNDIKIVHHDG-PFKIMQRLLREYKRFKEHIISITLRTTGGI : 291
OpMNPV  : EYFT-YNDVKIVHHEG-PFKIIQRLLREYKRFKEHIVSITLRTTGGV : 292
McNPV   : EYKVVKRQVPVRNYHD-LYATVKSIIAE-KQPYFQISEINLRTYGRT : 205
SeMNPV  : EYTVTETRTEVCNYHN-FYKMVKSIIRE-KKTNFHITEIKLSTYGRD : 218
HzSNPV  : ERRVDIVRNETYKVND-LYATIQDIISN-KNPHFQITKISFRTIGRV : 206
HaSNPV  : ERRVDIVRNETYKVND-LYATIQDIISN-KNPHFQITKISFRTIGRV : 206
TnSNPV  : EHTCTTEEISVVGRHD-FFGTLKDIVDK-KKENFQITKILLRTYGRT : 237
LdMNPV  : EYTT--SVRNVRDETLDLARSFRDILAQ-KAPNHQIVRIALRTYAQR : 194
SpltNPV : PSRCSARTVRITDTKRDIMKFFDETLRK-KHDHFQIISIELQTYGGV : 162
XcGV    : EPGK-KKCRRIGLQPVDIPTVLDDLVQS-KSLCEDIESVTLRNANND : 188

AcMNPV  : VLRETYTNIVLQRYRSMCTRPDVVDDVNCFILQEPSEMMAALQDNRF : 343
BmNPV   : VLRETYTNIVLQRYRSMCTRPDAADDVNCFILQEPSEMMAALQDNRF : 344
CfMNPV  : ALKVINDNVRLMRYRNIYKEPTASDDVNCFTVSSRSTLMETIDNGTF : 338
OpMNPV  : VLKEINHNVRLMRYRNIYKEPTASDDVNCFTVSSSSELMKALDSGAF : 339
McNPV   : LFTDTDFDCTIKST--FQDNPDVPSAEFEFYPVESS--MLKVVKTFD : 248
SeMNPV  : LPIESDDNCAMRN---GQNEYNEHVYELEFHPSPST--MIPFVQTYA : 260
HzSNPV  : FFDETFEDMFVEK-----------RGTISVVPGPSK--MLEFLSKPF : 240
HaSNPV  : FFDETFEDMFVEK-----------RGTISVVPGPSK--MLEFLSKPF : 240
TnSNPV  : LFTESDRNCIIKTT---SSKIGCDDNTFQLYFGDSK--MMDFFASMG : 279
LdMNPV  : LFAEELQNMYYDA---------EAPDELSVAPARSA--MLDFRKTHT : 230
SpltNPV : VFTERFEDAVMER------------NLIRILPNGSR--LAQVIESDF : 195
XcGV    : EVVARDDIYYLRV------------DWGSDITFSVP---TQFTRQLL : 220

AcMNPV  : SGIK-GTVFATVKVKKFTQVLDG---AITFPLKPTTNNYCKLCKKTK : 386
BmNPV   : SGIK-GTVFATVKVKKETEVLDGG--AITFPSRPTINNYCKLCKKTK : 388
CfMNPV  : NAIQ-GTVFAEIYGFAIQKFVTG---VVTFPLKPVKGGHCISCKKNK : 381
OpMNPV  : DSIQ-GTVFAEIYGFAIQKFVTG---VITFPVRPVKGNYCAVCKKNK : 382
McNPV   : DRTL-RDYFYEVTTRVYNTRFDY---VLYYNIKCTL--CCTFCTDAK : 289
```


https://etd.uwc.ac.za/

```
SeMNPV : EKKP-LTYFYTVTKRVYLSFFDY---AVCFPIKCIN--YCKMCKQDK : 301
HzSNPV : DFTPNFTYYYHVHVAVGREKQRY---VMYLEIPCLR--YCKLCTLEK : 282
HaSNPV : DFTPNFTYYYHVHVAVGREKQRY---VMYLEIPCLR--YCKLCTLEK : 282
TnSNPV : EQKL-LTYFYSVEKKVYKTTFNF---VLYFPIPCKR--FCILCTRHK : 320
LdMNPV : FINL--TYFYEIEKRVYHSVGHAAGYVAYFARPYAPLRSRSACIRCK : 275
SpltNPV: KFEN-KTYIYEIKMLVHHDPQTY---QVFVDRECGY--YCKKCTIKK : 236
XcGV   : NKSG--DYYLEIITRDYKEFQPF---YVFFHHGNNT--SCVACINKL : 260
                                                ▲  ▲

AcMNPV : LYYKN-----PVLYCTKCGFTNVYHFP-------EYSKFMYYFEAIK : 421
BmNPV  : LYYKN-----PVLYCTKCGFTNVYHFP-------EYSKLVYYSKAIK : 423
CfMNPV : MYYTN-----PIINCSKCGFTNRYIFKN------KYDHIYFHAEAVQ : 417
OpMNPV : MYYSN-----PVLCCSKCGFTNRYIFNG------KYDDIYFHPEAVQ : 418
McNPV  : LYKKTH----PIMYCTKCGFTDALKFRD----SNIINGIKYYKQCVV : 328
SeMNPV : LYFKIN----PVLYCSKCGFTDASFFKN----APFLKNTVFLAKCVK : 340
HzSNPV : QHKGY-----PVVWCSVCGYTDTMYYDE----EFLHFQNMEYESFRL : 320
HaSNPV : TT--------------------------------------------- : 284
TnSNPV : MYLKKH----IVLYCSQCGFTDAMFFTRN---KLDVSALKFYSVCVK : 360
LdMNPV : SRFYKNN---PILYCSRCGFMNRIYFKPSQRNELDYPSLVYLQRCVQ : 319
SpltNPV: QYRGN-----PIMFCSRCGNTDSKYNQQ-----FVMDNIGYDKKRLK : 273
XcGV   : CLTQKKKRLVPVLFCNNCGFTDPNYWSS---TTPNVYPFWLDSYDYK : 304
                    ▲  ▲

AcMNPV : SFEMH---NEMIIYYDLKMYKKLINIVNNNV---------------- : 449
BmNPV  : NFEMH---NEMIIYYDLKMYKKLINIVNNNV---------------- : 451
CfMNPV : THAIN---GEFIRYYDLKLHAKICRERLENYEIE------------- : 448
OpMNPV : THATHSTHGEFVRYYNLKLHAKICRERLEEYEAQNLQ---------- : 455
McNPV  : EKTKN---LKCIRYYDLDLYKSLVKKNKK------------------ : 354
SeMNPV : AKTLP---PTRIYYYDMNLYKSYMNKASLSSSPTASKNKLFDFKPKA : 384
HzSNPV : RPMYNKKKTECIIYYKLPFMPPSFLKNKTQSTLLSVTKQ-------- : 359
HaSNPV : ----------------------------------------------- :   -
TnSNPV : VKTIK---PKRIYYYDMNLYKKINV---------------------- : 382
LdMNPV : AVKTA---SYCIIYYDLNMYKRHKVK--------------------- : 342
SpltNPV: WKIGK---KYTLLYYDISDCIDKVLMPDYVM---------------- : 301
XcGV   : RTYWKCRKKVNLMLYDVDVSI-------------------------- : 325

AcMNPV : ------ :   -
BmNPV  : ------ :   -
CfMNPV : ------ :   -
OpMNPV : ------ :   -
McNPV  : ------ :   -
SeMNPV : LKFGKI : 390
HzSNPV : ------ :   -
HaSNPV : ------ :   -
TnSNPV : ------ :   -
LdMNPV : ------ :   -
SpltNPV: ------ :   -
XcGV   : ------ :   -
```

UNIVERSITY of the WESTERN CAPE

Figure 4.3 Alignment of deduced amino acid sequences of twelve baculovirus *me-53* genes. Two pairs of the completely conserved cysteine residues at the N-termini and the conserved cysteine residues of putative zinc finger-like motifs at the C-termini are

https://etd.uwc.ac.za/

indicated by open and solid triangles, respectively. The basic amino acids in the zinc finger-like domains are in bold italic. The sequences used in this study (with GenBank accession number) were: AcMNPV, *Autographa califonica* MNPV (NC001623); BmNPV, *Bombyx mori* NPV (NC001962); CfMNPV, *Choristoneura fumiferana* MNPV (unpublished); OpMNPV, *Orgyia pseudotsugata* MNPV (NC001875); McNPV, *Mamestra configurata* (AF467808); SeMNPV, *Spodoptera exigua* MNPV (NC002169); HzSNPV, *Helicoverpa zea* SNPV (NC003349); HaSNPV, *Helicoverpa armigera* SNPV (NC003094); TnSNPV, *Trichoplusia ni* SNPV (unpublished); LdMNPV, *Lymantria dispar* MNPV (NC001973); SpltNPV, *Spodoptera litura* NPV (NC003102); XcGV, *Xestia c-nigrum* GV (NC002331).

The motif patterns of the seven compared ME-53s are either the same as, or similar to that found in the AcMNPV ME-53 (Fig. 4.4). However, unlike the zinc finger motifs of the above ME-53s consisting only of clustered cysteine residues, the motifs in the McNPV and HzSNPV ME-53s are containing a histidine residue, respectively, but two in the TnSNPV ME-53, while the HaSNPV ME-53 only has three cysteine residues clustered in the region. The other two completely conserved cysteine pairs are clustered at the N-termini of the ME-53s (Fig. 4.3). The SWISSPROT database search showed that there was no strong sequence homology with any other known classes of zinc finger motifs in all these ME-53s.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

AcMNPV	$C-X_2-C-X_{13}-C-X_2-C$
BmNPV	$C-X_2-C-X_{13}-C-X_2-C$
CfMNPV	$C-X_2-C-X_{13}-C-X_2-C$
OpMNPV	$C-X_2-C-X_{13}-C-X_2-C$
LdMNPV	$C-X_2-C-X_{12}-C-X_2-C$
SeMNPV	$C-X_3-C-X_2-C-X_{14}-C-X_2-C$
SpltNPV	$C-X_3-C-X_2-C-X_{13}-C-X_2-C$
XcGV	$C-X_2-C-X_4-C-X_{13}-C-X_2-C$
HaSNPV	$C-X_3-C-X_2-C$
McNPV	$C-X_3-C-X_2-C-X_9$-His-X_4-C-X_2-C
HzSNPV	$C-X_3-C-X_2-C-X_5$-His-X_7-C-X_2-C
TnSNPV	$C-X_3-C-X_2-C-X_2$-His-X_6-His- X_4-C-X_2-C

Figure 4.4 Patterns of the zinc finger-like motifs in the baculovirus ME-53s. The identical and similar patterns are grouped, respectively. The histidine residues are present in the motifs of the bottom three baculoviruses ME-53s but not in the top nine.

In the TnSNPV *me-53* promoter region, three TATA box-like sequences are located at the position of -64, -101 and -162 nucleotide (nt), respectively (Fig. 4.2), and no consensus early promoter motif [ATCA(G/T)T] (Cherbas, 1993) are found. However, a putative early transcription start site CGTGC (Lu and Carstens, 1991; Carstens *et al.*, 1993) and a late transcription initiator element TAAG are located at the positions of -24 and -32 nt, respectively. Several GC and CGT motifs are also found in the regulatory region (Dickson and Friesen 1991), A polyadenylation signal sequence is identified 41 nt downstream of the TAA stop codon.

4.4.2 Infection of the recombinant viruses and GFP expression

Compared with the uninfected cells (Fig. 4.5A and D), the symptoms of infected cells at 48 h p.i. are clearly observed in Fig. 4.5B and E . The infected cells are swollen and their

https://etd.uwc.ac.za/

nuclei are enlarged. The space of cytoplasm of the cells are squeezed due to enlarged nucleus.

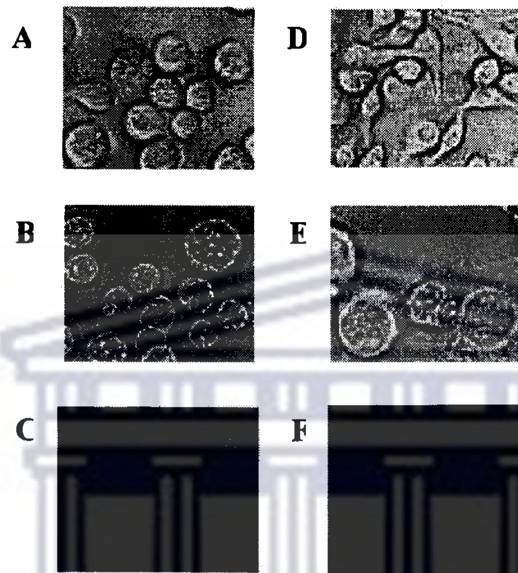


Figure 4.5 Uninfected and AcBacTnme53GFP-infected Sf9 (left column and High5 (right column) cells. The uninfected and infected Sf9 cells were observed in bright field (A and B). the infected cells were examined under fluorescence microscopy (C). High5 cells are following the same order: uninfected, infected, infected with fluorescence (D, E and F).

To assess whether the TnSNPV *me-53* gene was introduced into cells by viral infection, expression of GFP in the infected cells was an indicator. However, the subcellular localization of the TnSPNV ME-53 protein cannot be determined as it was not fused with GFP but homogeneous fluorescence from GFP expression can be observed in the

https://etd.uwc.ac.za/

cytoplasm and nuclei (Fig. 4.5C and F). It appeared that the intensity of fluorescence was stronger in nuclei than in cytoplasm and also in Sf9 than in high5 cells.

4.4.3 Transcriptional analyses of the TnSNPV and authentic *me-53*s of AcBacTnme53GFP

The 5' end mappings of the TnSNPV *me-53* transcripts indicated that the early (3 h p .i.) and late (24 h p. i.) transcription start sites of the gene, in each case, were identical in the two different cell lines (High5 and Sf9). The most distal start site of the *me-53* at 1, 3, 6, 9 and 12 h p.i. was mapped at -34 in infected High5 cells, while the most proximal start site was at -31 at 18, 24, 36 and 48 h p.i., in the same infected cells (Fig. 4.2). In each case at least three clones were sequenced for analyses. The early transcript initiated at the C, two bases immediate upstream of a conserved late transcription motif TAAG, while the late one at the first A of the TAAG motif (Fig. 4.2). The early and late transcription start sites of the authentic *me-53* gene of AcBacTnme53GFP, in each case, were mapped identical in both infected cell lines at the corresponding time points (3 and 24 h p.i.). The results are identical to those of the wild type AcMNPV *me-53* (Fig. 4.7) (Mans and Knebel, 1997). In addition, the corresponding results of the 5' end mappings of the authentic *me-53* gene and the TnSNPV *me-53* suggested that both genes were transcribed simultaneously.

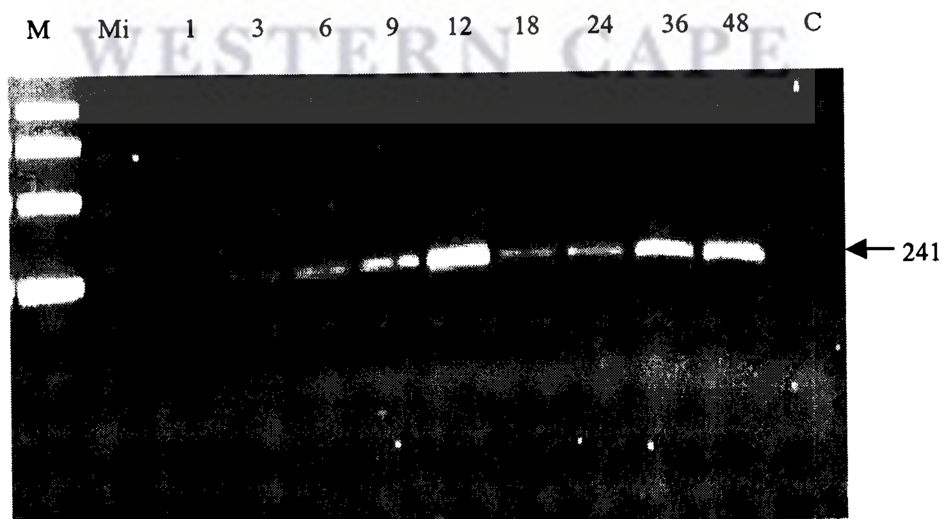
The sequencing results of the 3' RACE of the TnSNPV *me-53* transcripts from the High5 and Sf9 cell lines showed that a poly (A) chain attachment site of the TnSNPV *me-53* was located 21 nt downstream of a polyadenylation signal AATAAA at the last C in the

https://etd.uwc.ac.za/

sequence CTAC (Fig. 4.2). The sizes of the early and the late transcripts are 1244 nt and 1241 nt, respectively, excluding its poly (A) tail.

RT-PCR analysis (Fig. 4.6) was performed in order to obtain information about the temporal expression of the TnSNPV *me-53* transcripts. A single fragment of 241 bp was amplified at each time point using the gene specific primers Tn241F and Tn241R. No PCR products were obtained from the mock and AcBacGFP-infected samples when the same primers were used. The analysis revealed that the transcription of the TnSNPV *me-53* can be detected at 1 h p.i. and remain detectable until 48 h p.i. The increase of transcription level occurred in two different temporal periods: one from 1 h p.i. to 12 h p.i. and other from 18 h p.i. to 48 h p.i. From this analysis, combined with the 5' RACE results of the TnSNPV *me-53* obtained from infected High5 cells at the different p.i. times, it was concluded that the early transcript increased gradually to a peak level from 1 h p.i. until 12 h p.i., while the late one started from a low level at 18 h p.i. to a higher level at 48 h p.i.

UNIVERSITY of the
WESTERN CAPE



https://etd.uwc.ac.za/

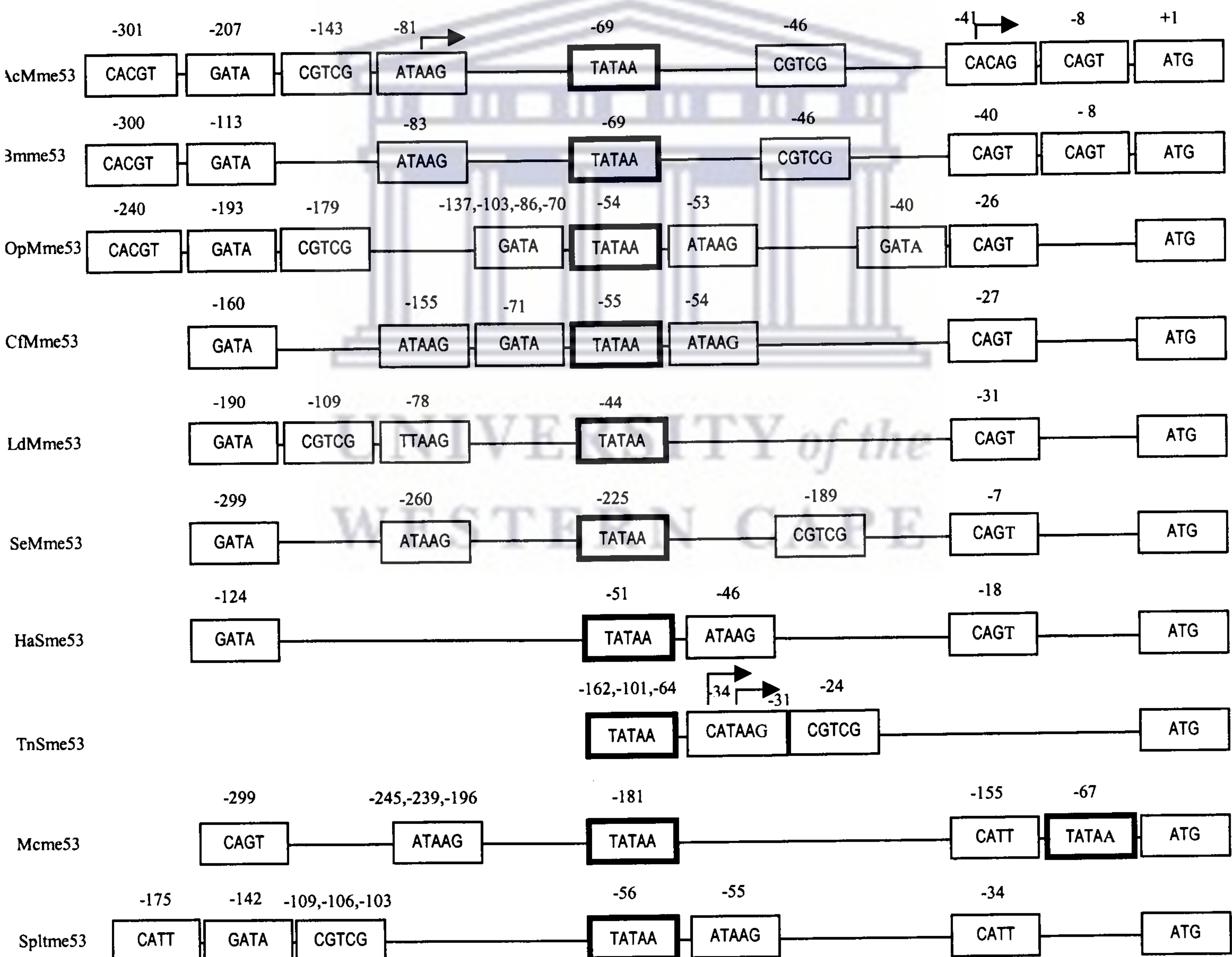
Figure 4.6 Temporal expression of the TnSNPV *me-53* transcripts by RT-PCR analysis. Times p.i. are indicated above the lanes. M and Mi represent the marker and the mock-infected sample, respectively. The letter C indicates a negative control: the AcBacGFP-infected sample at 12 h p.i. The results of PCR without RT on RNA templates at the above corresponding time points are not shown. The sizes of amplicons are shown in bp.

4.4.4 Analysis of the 5' non-coding regions of the NPV *me-53*s

Unlike NPV *ie-1* gene promoters, the DNA sequences of ten NPV *me-53* promoters are not conserved. To obtain detailed information on the promoter regions of *me-53* genes, existing regulatory elements of 10 baculovirus *me-53*s were grouped and highlighted (Fig. 4.7). A putative early initiation motif CAGT or CATT is conserved at the different positions upstream of the translation start sites of these *me-53*s but not present in the TnSNPV *me-53*'s. Two early motifs CAGT were present in the AcMNPV and BmNPV *me-53* promoters, while two CATT sequences were also found in this region of the SpltNPV *me-53*. The 5' non-coding region of the McNPV *me-53* contains a CAGT and a CATT sequence. The other putative early transcription start motif CGTGC suggested by Lu and Carstens *et al.* (1991, 1993) was located in the upstream regulatory regions of seven *me-53*s, excluding those of the CfMNPV, HaSNPV and McNPV *me-53*s. A TATA box-like motif was conserved in all NPV *me-53* promoter regions but no TATA motif observed in the region of XcGV *me-53* (data not shown). Two TATA-like sequences were found in the promoter region of the McNPV *me-53* and three in the region of the TnSNPV *me-53*. A putative late conserved transcription element ATAAG was located at either upstream or downstream of a TATA box in the promoter regions of most *me-53*s.

https://etd.uwc.ac.za/

Two ATAAG sequences were found in the regulatory region of the CfMNPV and TnSNPV, and three in the McNPV *me-53*, respectively. The six host transcription factor-binding-like motifs GATA (Kogan and Blissard, 1994) were present at different positions of the upstream regions of OpMNPV *me-53* relative to translation start codon, two in the CfMNPV *me-53* and one in each of six NPV *me-53*s but not observed in the regions of the McNPV and TnSNPV *me-53*s. Another host transcription factor-binding-like sequence CACGTG (Kogan and Blissard, 1994) is present in the upstream regions, relative to the translation start codon, of the AcMNPV, BmNPV and CfMNPV only.



https://etd.uwc.ac.za/

Figure 4.7 Comparison of regulatory elements within the upstream regions of ten NPV *me-53*s. The early and late transcription start sites of the AcMNPV (the wild type AcMNPV *me-53* and authentic *me-53* of AcBacTnme53GFP) and the TnSNPV *me-53*s are indicated by arrows. A single number, or multiple numbers, over each box represent positions of the putative regulatory motifs of these genes. The order of positions for all the elements of each gene is from proximity to distance relative to putative ATG translation start codon. Although the position of each element relative to its compared counterparts may vary, the same elements among different genes are aligned up except a TATA-like sequence at –67 nt upstream of the translation start codon ATG of the McNPV *me-53*. The TATA-like sequences are highlighted by the open boxes with thick edges.

4.4.5 Comparative phylogenetic analysis of IE-1 and ME-53 proteins

An alignment of twelve baculovirus ME-53 amino acid sequences is presented in Fig. 4.3. The predicted amino acid sequence identities between the ME-53s of group I and 2 baculoviruses range from 11 to 19% (Table 4.1). The highest ME-53 amino acid identity is 91% between AcMNPV and BmNPBV, followed by 77% between CfMNPV and OpMNPV and between HzSNPV and HaSNPV, respectively. The TnSNPV ME-53 shares the highest amino acid homology with both McNPV and SeMNPV (33% identity and 52% similarity) homologues, which were previously identified as the members of group II-A, and lower homologies with the rest of Me-53s (Table 4.2).

https://etd.uwc.ac.za/

	BmNPV (%)	CfMNPV (%)	OpMNPV (%)	McNPV (%)	SeMNPV (%)	HzSNPV (%)	HaSNPV (%)	TnSNPV (%)	LdMNPV (%)	SpltNPV (%)	XcGV (%)
AcMNPV	91	37	38	16	18	15	12	18	16	14	11
BmNPV		36	39	16	18	15	13	19	17	14	11
CfMNPV			77	17	18	13	11	17	15	14	13
OpMNPV				16	18	15	12	17	16	14	12
McNPV					**39**	27	24	**33**	27	22	15
SeMNPV						24	21	**33**	26	20	16
HzSNPV							77	24	23	24	15
HaSNPV								22	20	21	13
TnSNPV									27	18	15
LdMNPV										20	15
SpltNPV											18

Table 4.1 Identity values of the amino acid sequences of twelve baculovirus ME-53s. The baculovirus group I includes AcMNPV, BmNPV, CfMNPV and OpMNPV and the rest of NPVs belong to group II. The numbers represent percentages of sequence identity among the ME-53s. The scores shared by the TnSNPV, SeMNPV and McNPV ME-53s are in bold.

	AcMNPV (%)	BmNPV (%)	CfMNPV (%)	OpMNPV (%)	McNPV (%)	SeMNPV (%)	HzSNPV (%)	HaSNPV (%)	LdMNPV (%)	SpltNPV (%)	XcGV (%)
Identity	18	19	17	17	33	33	24	22	27	18	15
Similarity	36	37	32	32	**52**	**52**	41	37	42	36	30

UNIVERSITY of the
WESTERN CAPE

Table 4.2 Amino acid sequences of eleven NPV IE-1s compared with the TnSNPV IE-1. The greatest homologies shared between either SeMNPV or McNPV and TnSNPV, IE-1s are in bold.

The patterns of the homology comparisons among the ME-53 homologues are similar to those of the baculovirus IE-1s (Chapter 2). The predicted amino acid sequence identities among the eleven NPV IE-1s of group I and II baculoviruses range from 14 to 21%. The

https://etd.uwc.ac.za/

highest scores of sequence identities of the IE-1s are obtained between AcMNPV and BmNPV (96%), between CfMNPV and OpMNPV (73%) and between HzSNPV and HaSNPV (96%). The TnSNPV IE-1 shares the highest amino acid homology with the group II-A SeMNPV IE-1 (22% identity and 40% similarity) and lower homologies with other group I and II baculovirus IE-1s (Chapter 2).

To analyze phylogenetic relationships among baculovirus ME-53s a rooted phylogenetic tree was constructed using neighbor-joining method. The tree was divided into two main branches, group I (AcMNPV, BmNPV, CfMNPV and OpMNPV) and II (rest of NPVs) NPVs (Fig. 4.8). The group II NPVs were further clustered into two subgroups, II-A (SeMNPV and McNPV) and II-B (HzSNPV, HaSNPV and SpltNPV) (Bulach et al., 1999). The TnSNPV ME-53 was included in subgroup II-A with high bootstrap value. The topology of the tree base on the data of the compared ME-53 proteins is consistent with that of the tree constructed with the IE-1 data, except LdMNPV ME-53 which was clustered into the different subgroup in the different tree case (Fig. 4.8). The inclusion of the LdMNPV ME-53 in subgroup II-A was supported by a comparatively high bootstrap score in tree A, whereas its inclusion in subgroup II-B was with a low bootstrap value in tree B. The low bootstrap value was obtained for the separation of two subgroups in both tree cases, but the tree (B) for ME-53s was supported by a bootstrap value over 50%, while the tree (A) for IE-1s was below 50%. This also occurred for inclusion of the SpltNPV in subgroup II-B in both tree B (bootstrap score: 53.8) and tree A (bootstrap score: 48.7).

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

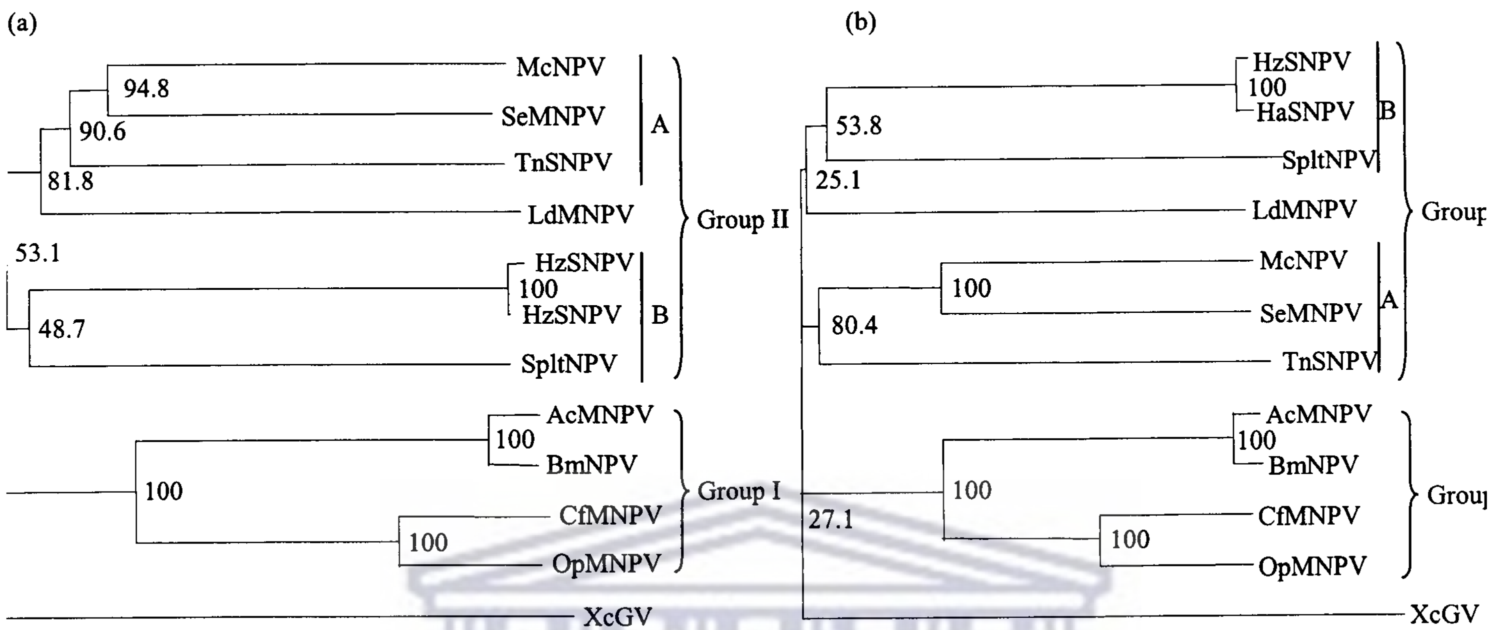


Figure 4.8 Rooted neighbor-joining trees based on alignments of deduced baculovirus ME-53 (a) and IE-1 (b) proteins. Numbers at each node represent percentages of bootstrap values (1000 trials). Branch lengths are proportional to relatedness. The letter A and B indicate subgroup II-A and II-B, respectively. XcGV taxa was used as an outgroup.

4.5 Discussion

4.5.1 Amino acid sequence analysis of the TnSNPV ME-53

The baculovirus major early-transcribed gene, *me-53* is postulated to be involved in regulation of gene expression on the basis of the presence of a zinc finger-like motif on its amino acid sequence. The zinc-finger motif was previously recognized as a DNA binding motif in *Xenopus laevis* transcription factor TFIIIA and other proteins (Miller *et al.*, 1985; Green *et al.*, 1988). The similar motif with clustered cysteine residues has also been found in the AcMNPV PE-38 (Krappa and Knebel, 1991) that was identified as a

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

baculovirus regulatory gene (Lu and Carsten, 1993). It was previously reported that deletion of a zinc finger-like cysteine cluster or replacement of the cysteine residues from a DNA-binding protein demolished sequence-specific DNA-binding activity (Green *et al.*, 1988; Kumar and Chambon, 1988). A comparison among the twelve baculovirus ME-53 amino acid sequences reveals that four or five cysteine residues clustered at the C-terminus are conserved among those proteins, excluding the HaSNPV ME-53 in which only three cysteine residues are clustered in that region (Fig. 4.3). However, the zinc finger-like motifs of the ME-53s showed no strong sequence homology with any known zinc finger domain in the SWISSPROT database. Further analysis of the zinc finger-like motif found that the patterns of the motifs in eight of twelve Me-53s (Fig. 4.4) are precisely identical or similar to the sequence of the form Cys-X_2-Cys-X_{13}-Cys-X_2-Cys which was previously identified as a potential metal-binding motif in the adenovirus E1A gene products that positively regulate the early promoters (Ricciardi *et al.*, 1981; Guilfoyle *et al.*, 1985). The other different feature of the motif, observed in three baculovirus ME-53s, is that one or two histidine residues exist in each of the zinc finger-like motifs (Fig. 4.4). The similar structures were found in the zinc finger-like domains of some of the proteins involved in mediated transcription regulation and protein-protein interactions (Freement, 1993; Gibson *et al.*, 1995). Although only three cysteine residues are clustered in the motif of the HaSNPV Me-53, the pattern of the sequence is the same as some parts of the zinc finger-like motifs in the other five ME-53s compared in this study (Fig. 4.4). Interestingly, the zinc finger-like domain-containing region of the ME-53s are rich in basic amino acids, specially rich in Lysine and arginine (Fig. 4.3). This type of structure was thought to be important for specific DNA binding (Kumar and

https://etd.uwc.ac.za/

Chambon, 1988; Weinberger *et al.*, 1985). In addition, two pairs of completely conserved cysteine residues, separated with two amino acids, were found at the N-terminus of each ME-53 (Fig. 4.3). There are 26 to 45 amino acids intervened between the two pairs of the cysteine clusters in the different ME-53s. The function of the cysteine clusters remains to be identified.

4.5.2 Accurate transcription initiation of the TnSNPV *me-53* promoter

To faithfully map the TnSNPV *me-53* in the context of AcMNPV genome, the BTI-Tn-5B1-4 cell (High5 cell) and Sf cell lines were selected for this study. This was based on two considerations: 1) Both cell lines are permissive to AcMNPV replication and have been used as the BEVS hosts (Wickham *et al.*, 1992; Wickham and Nemerow, 1993; Saarinen *et al.*, 1998); 2) The TnSNPV was originally isolated from diseased *Trichoplusia ni* larvae (cabbage looper) and High5 cells originated from the ovarian cells of Tn larvae (Granados 1994). Therefore, TnSNPV early promoters can be recognized and initiated accurately by the host factors of High5 cells. The previous *in vitro* studies (Mans *et al*, 1998; Knebel *et al.*, 1996) demonstrated that host RNA polymerase II in crude nuclear extracts, prepared from the uninfected insect cells, can precisely initiate the AcMNPV *me-53* promoter. Although transcription of the TnSNPV *me-53* occurred in the context of AcMNPV genome, the faithful and independent transcription of the gene can be achieved due to the fact that the expression of the baculovirus early gene depends on host RNA polymerase II and does not require any viral products. On the other hand, the early and late 5' ends mapping assays in this study indicated that the authentic *me-53* gene of AcBacTnme53GFP was faithfully transcribed in both cell lines, and the identical

https://etd.uwc.ac.za/

early and late transcription initiations of the TnSNPV *me-53* also occurred in these different types of infected insect cells. This suggested that the transcription initiations of the authentic *me-53* and the TnSNPV *me-53* did not interfere with each other. The orientation and location of the TnSNPV *me-53* in the locus of the polyhedrin gene (Fig. 4.1) also suggested that it was unlikely that transcription of the TnSNPV *me-53* was interfered by other flanking genes. Previously, a similar work was conducted by Kneble *et al.* (1996). In their experiment an AcMNPV/ Me-53 recombinant contained two copies of the *me-53* promoters which directed the transcription of an authentic *me-53* gene and a LacZ gene (ME53-LacZ), respectively. The latter was located upstream of the polyhedrin gene. The structure of the AcMNPV/Me-53 recombinant is similar to that of AcBacTnme53GFP, except for the genes ME53-LacZ, GFP and the TnSNPV *me-53* (Fig. 4.1). The primer extension assay showed that the early and late transcription start sites in the two identical *me-53* promoters (authentic *me-53* and ME53-LacZ promoters) were mapped as accurately as those in the wild type AcMNPV *me-53*. This suggested that an accurate transcription of the TnSNPV *me-53* in AcBacTnme53GFP may occur as the ME53-LacZ did in the AcMNPV/Me-53 recombinant. Furthermore, the baculovirus expression system we used has been widely used to express many heterologous genes (Saarinen *et al.*, 1998). Taken together, the above evidence strengthened the fact that the TnSNPV *me-53* in the context of AcMNPV genome can be faithfully transcribed.

4.5.3 Early transcriptional analysis of the TnSNPV *me-53*

In the baculovirus early transcription, except in the cases of a CAGT motif serving as a transcription start site, some early genes lack the CAGT motif (Nissen and Friesen, 1989;

https://etd.uwc.ac.za/

Crawford and Miller 1988; Lu and Carstens, 1991). In these cases the TATA motif may play a key role in directing proper downstream initiation. A previous study investigating the role of the TATA motif in directing the transcription of the AcMNPV *p35* promoter which lacks a CAGT motif (Dickson and Friesen, 1991) suggested that the TATA elements (possibly in combination with downstream elements) constituted the basal promoter of the gene and were responsible for establishing the position of the RNA start site. It was demonstrated that upon insertion of DNA between the TATA box and the normal RNA start site, the distance to the utilized RNA start site remained constant (29 to 34 bp). Furthermore, the GC and CGT motifs in the upstream region of the TATA box were exhibited to be distant regulatory elements for stimulation of the basal transcription of the AcMNPV *p35* promoter (Dickson and Friesen, 1991). By inspection of the promoter region of the TnSNPV *me-53*, no CAGT motif was found and only a putative early CGTGC motif was present in this region. However, the C in the sequence TCATAAG upstream of the CGTGC sequence was identified as the early transcription start site of the gene in the 5' RACE but not the CGTGC sequence (Fig. 4.2). The distance from the early functional initiation site of the TnSNPV *me-53* to its TATA box is 30 bp, which is precisely the same as that of the early initiation site to the TATA box in the case of the AcMNPV *p35*. In accordance with the position of the early initiation site of the TnSNPV *me-53* relative to the TATA box and several copies of GC- or CGT-like motifs presented in the upstream region of the TATA box (Fig. 4.2), it is possible that the TATA motif within the TnSNPV *me-53* promoter may play a similar functional role to that demonstrated in the case of the AcMNPV *p35*.

https://etd.uwc.ac.za/

To investigate temporal transcriptional pattern of the TnSNPV *me-53*, the 5' RACE and RT-PCR at the different p.i. times were performed. The results indicated that a detectable level of the early transcription occurred from 1h p.i and dropped from a peak after 12 h p.i. Although the early and late transcripts of the TnSNPV *me-53* could not be distinguished by their sizes in the RT-PCR assay, two peak levels of the transcripts observed at 12h p.i. and 48h p.i. (Fig. 4.6), together with the transcriptional results of the 5' RACE, indicated that it was more likely that a switch from early to late transcription occurred after 12 h p.i. In accordance with transcription of strictly defined early genes, the early transcription usually reaches a peak between 6 and 12 h p.i. and declines thereafter when late viral transcription is vigorous (Friesen, 1997). It appeared that the temporal early transcriptional pattern in the cases of the TnSNPV *me-53* is consistent with the suggested definition above. However, the pattern of the early transcript of the TnSNPV *me-53* differed from that of the wild type AcMNPV *me-53* analyzed previously by Mans (1998) and Knebel *et al.* (1996). This pattern difference also occurred in a similar study conducted by Knebel *et al.* (1996) but it did not influence accurate transcription of the early gene in their experiments. A possible explanation for the change of the transcriptional pattern could be due to the impact of either modified regulation or difference of mRNA stability on the transcription (Knebel *et al.*, 1996). The TnSNPV *me-53* may also be a similar case in this regard.

https://etd.uwc.ac.za/

4.5.4 Early regulatory elements of the baculovirus *me-53*s and strategy for determination of the functional role of the TATA motifs in the TnSNPV *me-53* promoter region.

The transcription initiation of baculovirus early genes is known in connection with several promoter motifs. Among these a conserved sequence CA(G/T)T usually located upstream of the ATG translation start site and downstream of a TATA-box has served as a mRNA transcription start site in many cases (Blissard and Rohrmann,1989; Carson *et al.*, 1991; Krappa and Knebel, 1991; Guarino and Smith, 1990; Chisholm and Henner, 1988) but in some cases the CGTGC motif was identified to be an early transcription initiation site (Lu and Carstens, 1991; Carstens *et al*; 1993). A study showed that substitutions of CAGT sequence decreased the level of transcription significantly in the case of *gp64* (Blissard *et al.*, 1992). This, and the other studies, demonstrated that replacement of CAGT with TCTA did not alter the transcription start site relative to the TATA box, while deletion of the TATA box resulted in loss of all transcriptional activities (Blissard *et al.*, 1992; Blisard and Rohrmann, 1991). All this evidence suggested that the TATA box sequence was the primary element controlling the start site selection and the CAGT sequence did not appear to be essential for transcription in the presence of a TATA box but may be important for efficient basal transcription initiation (Blissard and Rohrmann, 1991; Blissard *et al.*, 1992). However, another experiment performed by Guarino and Smith (1992) showed that alternation of the proximal CAGT motif in the early *39k* promoter, which contains two functional TATA boxes, abolished transcription from the site and suggested that the CAGT motif functioned as an initiator element. Although the above conclusions about the relationship of the regulatory roles of TATA and CAGT motifs from the studies of the two different gene promoters (*gp64* and

https://etd.uwc.ac.za/

39k) may vary, the CAGT sequence indeed serves as a transcription start site in many of the early genes, such as *ie-1, 39k, gp64*. A comparison between the promoter regions of the ten baculovirus *me-53*s indicated that at least a TATA-like box and a CAGT or a CATT sequence are present in the regions, except for the TnSNPV *me-53* which lacks a CAGT or CATT sequence (Fig. 4.7). This observation appears to suggest that TATA and CAGT or CATT may function in a cooperative manner in the early transcription of these *me-53* genes as suggested in the case of *gp64* (Blissard *et al*., 1992). However, the regulatory roles of TATA and CAGT motifs in the case of the NPV *me-53*s need to be further investigated. On the other hand, in the case of the XcGV *me-53*, no TATA box was found and only a CAGT motif was present in the upstream region, relative to the putative translation start site of the gene (data not shown). A similar case was found in the gypsy promoter (*Drosophila* retrotransposon promoter) which also lacks a TATA element (Jarrell and Meselson 1991). It was reported that a CAGT motif in the gypsy promoter served as an initiator for the gypsy retrotransposon. Other previous studies demonstrated that transcription could initiate properly from a CAGT motif after either replacement or deletion of the TATA box in the case of *ie-1, 39k* and *gp64* (Pullen and Friesen, 1995b; Kogan *et al*., 1995). These studies suggested that host factors are sufficient for TATA-less initiator activity from the CAGT motif and indicated that CAGT as an initiator element is sufficient to determine the RNA start site in the absence of a TATA element. Although there is no TATA element presented in the promoter region of the XcGV *me-53*, the gene may be transcribed in a TATA-independent manner as suggested above.

https://etd.uwc.ac.za/

Another unusual feature of the TnSNPV *me-53* is that three TATA-like sequences are present in its upstream regulatory region. It was shown in the previous experiment that dual TATA boxes in the early *39k* gene promoter could direct transcription from two functional early initiation sites independently by different mechanisms (Guarino and Smith, 1992). In the case of the TnSNPV *me-53* there is no evidence showing the occurrence of multiple early transcription, and therefore this may not be the case of the TnSNPV *me-53*. It is possible that some of the TATA sequences presented in the regulatory region of the TnSNPV *me-53* may not function as regulatory elements or function together in a cooperative manner. To determine the possible functional role of three TATA motifs in the upstream region of the TnSNPV *me-53*, site direct mutagenesis and 5' deletion of the promoter region can be introduced into the investigation. The strategy for point mutation of the *me-53* promoter is illustrated below (Fig. 4.9).

	-162	-101	-64	-34	+1
1	TATAAA	TATAAA	TATAAA	TTCA*TAAG*	ATG--coding region or a CAT
2	TAAAAA	TAAAAA	TATAAA	TTCA*TAAG*	ATG--coding region or a CAT
3	TAAAAA	TATAAA	TAAAAA	TTCA*TAAG*	ATG--coding region or a CAT
4	TATAAA	TAAAAA	TAAAAA	TTCA*TAAG*	ATG--coding region or a CAT
5	TATAAA	TATAAA	TATAAA	TTTA*TAAG*	ATG--coding region or a CAT
6	TAAAAA	TAAAAA	TAAAAA	TTCA*TAAG*	ATG--coding region or a CAT

Figure 4.9 Point mutation of TATA motifs and the RNA start site of the TnSNPV me-53 promoter region. The five open boxes in each row represent the promoter and either coding region of the *me-53 gene* or a CAT gene. Three TATA motifs, the RNA start site and the *me-53* coding region or a CAT gene are indicated in each respective open box.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Numbers (1-6) on the left side of each row of open boxes represent the wild type and different mutants of the early *me-53* promoter region in different recombinant AcMNPVs, respectively. The different mutation points in the promoter region were underlined in bold. The early transcription initiation site of the *me-53* was mapped at C (in bold), which is located 2 nt upstream of a late transcription start site TAAG (in italic). The locations of the early regulatory elements in the wild type promoter region are indicated with numbers. The putative translation start codon ATG of *the me-53* is designated as + 1.

The effect of the mutant TATA motifs on the early *me-53* transcription can be tested by using a baculovirus expression system, in vitro transcription assay, using High5 nuclear extracts, and transfection of a plasmid carrying both the *me-53* gene and its mutant promoter region. However, the latter is not recommended because the *me-53* promoter is rather weak if transfection rate is not very good, and the mRNA of the *me-53* may not be sufficient to be detected in a primer extension assay or a CAT assay which is required to fuse a CAT gene with the various mutant promoter regions. The following is an example for identifying the possible functional role of these TATA mutants in early *me-53* transcription by using a baculovirus expression system. The brief procedure is as follows: 1) Cloning the *me-53* gene with its different mutant promoter regions, illustrated above, and a wild type promoter into transfer vectors, as controls; 2) generating the recombinants containing the *me-53* carrying the different mutant early promoter regions; 3) primer extension analysis of mRNAs. Each of the mutant early promoters in each recombinant from 2 to 4 (Fig. 4.7) has only one TATA element and therefore the

https://etd.uwc.ac.za/

transcription directed by an active TATA, if this is so, will be detected in a primer extension assay. The recombinant 5 will reveal whether the utilization of the transcription start site will be controlled by a TATA box, while the recombinant 6 will determine which transcription mechanism, TATA-dependent or independent, will be utilized by the promoter. If the expected results are not attained, further experiments may need to be designed, such as abolishing only one TATA motif each time, etc.

4.5.5 Late transcription analysis of the TnSNPV *me-53*

A previous study suggested that the late promoters of the OpMNPV early *gp64* gene in the context of AcMNPV genome could be recognized and also be initiated accurately at only basal level in the permissive cell lines of either AcMNPV or OpMNPV but not activated (Garrity *et al.*, 1997). It was also thought that if TAAG sequences in an appropriate context are present within the heterologous gene, transcription may initiate within the gene (O'Reilly *et al.*, 1992). These suggestions led to the investigation as to whether the late promoter of the TnSNPV *me-53* could be recognized and initiated by the AcMNPV-encoded RNA polymerase (Guarino *et al.*, 1998). The temporal RT-PCR (Fig. 4.6) and 5' RACE analyses revealed that the late transcription of the TnSNPV *me-53* initiated from a TAAG motif, was detected at 18 h p.i. and reached more abundant level at 48 h p.i. The increase of the late transcripts was proportional to the time of infection. The late transcriptional pattern of the TnSNPV *me-53* is consistent with that demonstrated in the ME-53-Lacz of the AcMNPV/*ME-53* recombinant (described in section 4.5.3), but differed from that which occurred in the wild type *me-53* (Knebel *et al.*, 1996). However, as confirmed by Knebel *et al* (1996), the pattern difference did not

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

influence faithful transcription either from early or late transcription start sites of the ME53-LacZ in the AcMNPV recombinant (Knebel *et al.*, 1996). Collectively, the above analysis suggested that AcMNPV-encoded RNA polymerase may not be virus-specific. This polymerase can recognize not only the late promoter of the authentic *me-53* of AcBacTnme53GFP, and faithfully initiate the late transcription of the gene, as demonstrated in this study (Fig. 4.7), but also direct the late transcription of the TnSNPV *me-53* simultaneously. It is also likely that the late transcription of the TnSNPV *me-53* may be initiated faithfully, as occurred in the case of the ME53-lacZ in the AcMNPV/ME53 recombinant.

4.5.6 Late transcription initiation element of the NPV *me-53*s

The late transcription of baculovirus genes usually initiates within or near the consensus sequence (A/G/ T)TAAG, which is referred to as a TAAG site (Garrity *et al.*, 1997). Some of early genes have a functional TAAG which can serve as a late transcription start site during the late phase of infection (Knebel *et al.*, 1993; Nissen and Friesen, 1989) but some only have nonfunctional TAAG sequences (Carson *et al.*, 1991; Krappa and Knebel, 1991). By the inspection of the NPV *me-53* upstream regions, relative to their translation start sites, most *me-53*s contain only one TAAG motif (Fig. 4.7), and therefore the late transcription of these genes may initiate from the same motif TAAG as in the case of the AcMNPV and TnSNPV *me-53*s. Unlike these *me-53*s, the CfMNPV *me-53* has two TAAG motifs present in the regulatory regions, and McNPV *me-53* has three. Whether these TAAG motifs may function as their *gp64* counterparts (Garrity *et al.*, 1997), or may not function, as in the cases of *ie-2* (Carson *et al.*, 1991) and *pe38* (Krappa

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

and Knebel, 1991), is unknown. A previous study demonstrated that replacement of the sequences adjacent to the nonfunctional TAAG motif can convert it to be recognized by the host and virus-encoded RNA polymerases (Mans and Knebel, 1998). This might be the result of either simultaneous recognition or competition for the same template by both RNA polymerases. The other evidence (Garrity *et al*, 1997) suggested that the immediate context of flanking sequences of the TAAG, and not position, was primarily responsible for the selection of functional TAAG sites. The proper context of surrounding sequence of the TnSNPV *me-53* may also be a factor for its recognition by the virus-encoded RNA polymerase. A possible approach for distinguishing differences between flanking sequences of the functional and the non-functional TAAG motifs may be to look at whether they are rich in AT (Mans and Knebel, 1998). The AT richness was observed in the immediate upstream flanking sequence of the late TAAG motif of the TnSNPV *me-53* (Fig. 4.2). The function of the TAAG motif serving as a late transcription start site was confirmed by the 5' RACE and RT-PCR.

UNIVERSITY of the
WESTERN CAPE

4.5.7 Host factor-binding sites of the NPV *me-53*s

For better understanding of the gene regulation of baculovirus *me-53*, the predicted host transcription factor-binding site elements among the regulatory regions of ten baculovirus *me-53*s were compared (Fig 4.7). A putative host factor binding site GATA (Kogan and blissard, 1994) is present in the upstream regulatory regions of five *me-53*s, while multiple GATA sequences are observed in the case of the CfMNPV and OpMNPV *me-53*s. No GATA motif was found in the TnSNPV and McNPV *me-53*s. Another putative host transcription factor-binding site CACGTG (Kogan and blissard, 1994) is not

https://etd.uwc.ac.za/

conserved among these *me-53*s. However, the AcMNPV, BmNPV and OpMNPV *me-53*s each contain a putative GATA and a CACGTG motifs. A previous study demonstrated that a point mutation in either the GATA or CACGTG or both reduced transcriptional activity from the *gp64* early promoter (Kogan and Blissard, 1994). However, as characterized previously, these specific host binding sites are not required for TATA-dependent basal promoter activity although the binding of host binding factors to the sites is important for the activation of transcription (Blissard *et al.*, 1992; Blissard and Rohrmann, 1991). Due to the lack of both motifs, the TnSNPV *me-53* may be an exception for this type of gene regulation. The question raised in this study is how the host transcription binding factors may function in the GATA- and CACGTG-less gene regulation. These have not yet been investigated.

4.5.8 Phylogenetic analysis of the TnSNPV ME-53

A previous study on baculovirus IE-1 phylogenies (Chapter 2) suggested that TnSNPV is a member of subgroup II-A within NPV group II. This study showed that the TnSNPV IE-1 shares the highest amino acid sequence homology with the SeMNPV IE-1, and lower sequence homologies with other IE-1s of group I and II baculoviruses. However, the phylogeny based on a single gene may produce inaccurate phylogenetic inference due to unequal rates of evolution or lack of a robust phylogenetic signal (Herniou et al., 2003) and may not accurately represent the relatedness of TnSNPV to other baculoviruses. In contrast, the analyses of combining genes may reduce sampling errors and allow phylogenies to converge toward correct solution with good support (Mitchell et al., 2000). Herniou et al. (2001) analyzed 63 genes common to the nine lepedoteren

https://etd.uwc.ac.za/

baculoviruses for studying baculovirus phylogenies. The early genes, *ie-1* and *me-53* were included in their study, while other early genes, *ie-2, pe38, p35, gp64* (*ie-2* and *gp 64* specific to group I NPVs) and *ie-0* (in the absence of CpGV), and late genes previously used for phylogenetic studies, such as *egt* (not in XcGV), *p26* (not in XcGV), *chitinase, cathepsin* (both not in PxGV), and *p10* (not in CpGV) were excluded. More recently, Herniou et al (2003) used 30 genes shared among 13 baculovirus to reconstruct baculovirus phylogenies, excluding *ie-1, me-53* and *polh*, etc., because they are only present in lepidoteren baculovirus but not in all the baculoviruses. This suggested that the crucial criterion of gene selection for phylogenetic study is conservation of genes in all the baculovirus genomes, but not gene categories. In the study of Herniou et al. (2003) the reconstructed baculovirus phylogenies showed 4 major groups: GVs, group I and II lepidopteran NPVs and the dipteran NPV (Chapter 1). The phylogenetic analysis based on 30 genes conserved among all the baculovirus genomes showed no single gene tree was the same, but the construction of the majority rule consensus tree underlying different tree topologies was based on the data sets of all the shared genes. This consensus tree suggested that the stronger the phylogenetic signals of individual genes were in one branch, the more reliable the phylogenetic inference obtained would be and, also, that combining data sets from individual genes would possibly reinforce the weak signals. The previous bootstrap analysis of the TnSNPV IE-1 phylogeny showed that the trees based on IE-1 data sets were not well supported by bootstrap values in the separation of group II-A and II-B NPVs. The *me-53* gene, as a well-conserved early gene among all the lepidopeteran baculoviruses and its comparatively conserved amino acid

https://etd.uwc.ac.za/

sequence, was thus explored to infer the TnSNPV phylogeny among twelve baculoviruses for further resolving the deep branching of the TnSNPV phylogenies.

The tree topologies, specific to branching of group I and II, are consistent with and similar to one another, based on the alignment of deduced IE-1 and ME-53 proteins. The ME-53s of subgroup II-A and II-B NPVs are clustered together (Fig. 4.8). The TnSNPV was included in subgroup II-A with the support of high bootstrap values and highest amino acid sequence homologies shared with other members of this subgroup in comparison to homologies shared between TnSNPV and other subgroup I and II NPV Me-53s (Table 2). A moderate bootstrap support (53.1) in tree A suggested that the TnSNPV Me-53 shared a recent common ancestor with other group II NPV Me-53s, which was not well resolved in the previous TnSNPV IE-1 phylogeny. However, the resolution, in the case of the LdMNPV phylogeny, was obscured because its relatedness to subgroup II-A or II-B conflicted with one another in the two trees, but it appeared that LdMNPV Me-53 is more closely related to the TnSNPV homologue with a good bootstrap support in the ME-53 tree, while the LdMNPV IE-1 clustered with subgroup II-B in the IE-1 tree was only supported by a low bootstrap score (below 50%). Therefore, the LdMNPV IE-1 phylogeny in this tree (IE-1) has a low reliability. Perhaps, more ie-1 and me-53 sequence data and/or different gene data sets may be needed for resolving this disagreement. In conclusion, although there are slight differences between the two trees, as discussed earlier, both trees consistently agreed with the previous inference of the baculovirus phylogenies in terms of the separation of group I and II and subgroup II-A and II-B (Bulach et al., 1999), as well as inclusion of the TnSNPV into subgroup II-A (Wang et al., 2001).

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

4.5.9 A potential model system for improving bio-insecticidal agents

Baculoviruses have been genetically engineered for improvement of their properties as bio-control agents for more than ten years. Although different approaches have been used in an attempt to improve virulence and killing speed of the viruses, the baculovirus recombinant carrying multiple insect-specific toxin genes under the control of viral early promoters has not yet been reported. One of objectives of this study is to test a baculovirus model system which allows to express multiple foreign genes at the early stage of viral infection. The fluorescence observed in the AcBacTnme53GFP-infected Sf9 and High5 cells (Fig. 4.5C and F) and transcriptional analysis of the TnSNPV *me-53* showed that both genes under the control of the early promoters (TnSNPV *me-53* promoter and human cytomegalovirus (CMV) immediate early promoter) were expressed. It was also indicated that the recombinant virus replication was not compromised by the two heterologous genes because of successful amplification of the virus stocks. This study suggests that the model system, multiple heterologous genes under the control of early promoters, can be used to introduce multiple insect-specific toxin genes into a baculovirus polyhedrin locus for improving the properties of current bio-control agents.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Chapter 5

UNIVERSITY of the
WESTERN CAPE

Summary and Concluding Remarks

https://etd.uwc.ac.za/

Summary and Concluding Remarks

Baculoviruses are attractive biological alternatives to chemical insecticides of agriculture pests. Many attempts have been made to investigate the different characteristics of baculoviruses such as viral structure, mode of entry, replication, gene regulation and phylogeny. This research has led to a better understanding of this virus family. The application of this virus research basically falls into two main parts: one is genetic engineering of baculoviruses as bio-insecticidal agents, the other overexpresses foreign genes using the baculovirus expression system for the pharmaceutical and other research interests. These applications are established on the basis of understanding the fundamental issues of the baculovirus infection cycle. One of the important aspects is gene regulation which is highly complex and involves sequential and coordinated expression of early, late and very late genes. The virus regulatory cascade starts from early gene expression. The baculovirus replicative success relies on appropriate early gene expression and regulation. The immediate early gene 1 as a principle transregulator plays a crucial regulatory role in the successful viral replication.

UNIVERSITY of the
WESTERN CAPE

Until now only a few *ie-1s* have been identified and characterized. The characterization of the TnSNPV *ie-1* has contributed to a further understanding of the regulatory role of this gene in viral replication. The study revealed that early regulatory elements, CAGT and TATA, were almost completely conserved among the promoter regions of the eleven NPV *ie-1*s. This finding suggests that early transcription of NPV *ie-1s* may be mediated in a TATA dependent fashion and similar to one another. At the protein level, N-terminal

https://etd.uwc.ac.za/

regions consisted of the first third of each of IE-1 proteins are poorly conserved, while the C-terminal regions consisted of the remaining two-third well conserved. For the first time the acid-base profiles of eleven IE-1 proteins have been constructed. The comparisons of the acid-base profiles and charge distributions in the N-termini of IE-1s revealed that the first third portions of the proteins at N-termini are predominantly acidic and negatively charged, the characteristics of which have been found in a number of transactivators. Subsequent protein database searches indicated that the deduced TnSNPV IE-1 protein from 47-120 amino acids significantly matched the SET-domain of transcriptional regulators. The mapping result appeared to be similar to those determined by Rodems *et al* and Slack *et al* **(Chapter 2)**.

To identify the placement of the TnSNPV in baculovirus phylogeny previous attempts, using the late genes as a model, determined that the TnSNPV has belonged to group II NPVs. Owing to the small sizes of these late proteins, some of which have the highly conserved amino acid profiles such as *polh*, they were not as suitable for phylogenetic resolution as those with much longer and comparatively conserved amino acid sequences. More important, some of these late genes are not shared by all the lepidopteran baculoviruses and therefore do not reflect baculovirus evolution history. Thus far it has been found that the deduced TnSNPV IE-1 protein (739 aa) is the largest baculovirus IE-1 homolog in size, and has both poorly and highly conserved amino acid regions. The topologies of the two trees constructed, based on alignments of the entire IE-1s and the conserved regions of these proteins, were consistent with one another, and the trees indicated a high confidence in deep branching of the TnSNPV IE-1 phylogeny with the

https://etd.uwc.ac.za/

support of high bootstrap values (**Chapter 2**). The phylogenetic study not only supported the previous phylogenetic estimation of the TnSNPV as a member of group II NPVs, based on data sets of the late genes, but also further inclusion of the TnSNPV into the subgroup A within the group II.

Current evidence suggests that IE-1 transactivation is mediated in DNA-independent and *hr* enhancer-dependent manners. In the case of the former, no specific DNA sequences are involved in recognition by IE-1. For the latter, the binding of IE-1 to the 28 bp imperfect palindrome is required for stimulation of *hr*-mediated enhancement or DNA replication. Upon IE-1 synthesis it binds to the palindromic 28-mer as a dimer. The crucial step for the effective function of IE-1 is the proper interaction of oligomeric IE-1 with the 28 mer. The previous investigations of IE-1 protein-DNA binding have identified a b-HLH-like domain at the C-terminus of IE-1 involved in oligomerization and DNA binding, and suggested that the multiple residues and regions may also contribute to the DNA binding and oligomerization of IE-1. However, the direct binding region (s) has yet to be determined. The attempt to identify the binding region was made using AcMNPV IE-1 and a 60 bp repeat within the *hr*5 as a model system to replace the TnSNPV IE-1 and its binding DNA because of a lack of information of the TnSNPV *hr*s or other possible binding elements. The replacement of the system was based on the highly conserved structural similarity between both IE-1 proteins. The binding result indicated that the 130 aa region (Δ1-452; see Fig. 3.1) at the C-terminus bound to the 60 bp repeat directly and specifically. The finding provided the evidence for locating a direct DNA binding domain. It also suggested that a smaller region, such as 92 aa or a b-HLH-

https://etd.uwc.ac.za/

like domain fusion, can be used for future investigating the role of the b-HLH-like domain in the DNA-protein binding. In addition, the study provided the purified protein source for NMR or X-ray crystallographic analysis of the regulatory protein **(Chapter 3)**.

Apart from the *ie-1* gene another early gene was also identified in this study. The early major-transcribed gene *me-53* whose transcription, like *ie-1*, does not require any viral products, can be recognized and transcribed by host RNA polymerase II. It can be detected at the very early stage of post-infection. It was suggested that the gene could be a transactivator because of existence of a zinc finger-like motif within the C-terminus of the protein, the structure of which is believed to play a role in gene regulation. However, little is known about how the gene could be involved in gene regulation. The limited information of the gene is only obtained from AcMNPV *me-53*. No other *me-53* from other baculovirus species has been characterized before this investigation.

This study of the TnSNPV *me-53* has attempted to achieve two goals: 1). Identification and characterization of the gene; 2). Testing a model system for improving the properties of baculoviruses as bio-insecticides. For the first goal, some of the common and unique features of the gene have been investigated and presented **(Chapter 4)**. A comparative analysis between deduced protein sequences of baculovirus *me-53* genes was conducted. The highly conserved zinc finger-like motifs were found at the C-termini of the ME-53s. These motifs have similar structure patterns. This suggested that the early gene may be involved in gene regulation like others with similar structures. The TnSNPV ME-53 shares the highest homologies with SeMNPV and McNPV ME-53s that have been

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

suggested as the members of NPV subgroup II-A. The comparative phylogenetic study combined the data sets of IE-1s and Me-53s to further estimate TnSNPV phylogeny. The analysis based on the two data sets was in good agreement with the previous resolution of the TnSNPV IE-1 phylogeny and showed a high reliability supported by strong phylogenetic signals for the inclusion of the virus into NPV subgroup II-A (**Chapter 2 and 4**). At the transcriptional level, an early and a late transcription initiation site were identified in two infected cell lines. The early transcription initiation started from a unique early transcription site, two bases upstream of a late transcription start site TAAG. In addition, the occurrence of the late transcription of the gene suggested that the AcMNPV-encoded RNA polymerase may not be viral-specific. It can recognize the heterologous promoter in the case of the TnSNPV *me-53*. Finally, the analysis of 5' UTR of baculovirus *me-53*s was also conducted for exploring the possible transcription fashion of baculovirus *me-53* gene. The presence of conserved TATA-like boxes in the regulatory regions of NPV *me-53s* suggested that a TATA-dependent promoter activity may be essential for transcription of the early gene (**Chapter 4**). The above investigation provided valuable information for a further understanding of the functional roles of *me-53* in baculovirus regulation.

The second goal of this study was to develop a model system and provide useful information for the improvement of bio-control agents. Many research trials have been done in an attempt to improve the properties of baculoviruses as more virulent bio-insecticides with quick speed of action. In most trials deletion of a certain viral gene, insertion of an insect-specific toxin gene, or the combination of both approaches, were

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

applied to increase the killing speed of bio-insecticides to insects. Although a viral early promoter as well as late ones have been tried to achieve the above goal, insertions of multiple toxin genes, under the control of viral early promoters, have not yet been reported. The advantages of this approach are: 1) Multiple toxin genes, acting on the different physiological pathways of insect growth, can be expressed at the early stage of viral infection and therefore the infected insect may effectively cease feeding earlier; 2) Due to the foreign gene expressions directed by host factors, but not viral products, comparatively strong early promoters from different baculovirus species, or other heterologous systems, can be selected for this purpose, such as *ie-1*, *gp64* and CMV early promoters, so that the fast killing action of the genetically modified viruses may be achieved in the infected insect. The study has demonstrated that two early heterologous genes (GFP and the TnSNPV *me-53*) can function properly in the context of the AcMNPV genome without interfering with the virus infection and replication *in vitro*. The evidence from these two heterologous gene expressions in the AcMNPV recombinant suggested the model for improving bio-control agents by expressing multiple foreign genes at the early stage of viral infection in BEVS can be put into practice (**Chapter 4**).

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

References

Adlerz, W. C. (1971). Cabbage looper control on watermelon at Leesburg, 1968-1970. Florida state Horticulture Society proceedings: Annual Meeting. (8). 145-146.

Afonso, C. L., Tulman, E. R., Lu, Z., Balinsky, C. A., Moser, B. A., Becnel, J. J., Rock, D. L., and Kutish, G. F. (2001). Genome sequence of a baculovirus pathogenic for *Culex nigripalpus*. *Journal of Virology* 75, 11157-11165.

Ahrens, C. H., Leisy, D. J., and Rohrmann, G. F. (1996). Baculovirus DNA replication, p.855-875. In M. L. DePamphilis (ed.), DNA repilcation in eukaryotic cells. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.

Ahrens, C. H., Russel, R. L. Q., Funk, C. J., Enavs, J. T., Harwood, S. H., and Rohrmann, G. F. (1997). The sequence of the *Orgyia Pseudotsugata* multicapsid nucleopolyhedrovirus. *Virology* 229, 381-399.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215, 403-410.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Capped BLAST and PSI-BLAST: a new generaion of protein database search programs. *Nucleic Acids Research* 25, 3386-3402.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Apple, R. D., Bairoch, A., and Hochstrasser, D. F. (1994). Anew generation of information retrieval tools for biologists: the example of the ExPASy WWW server. *Trends in Biochemical Sciences* 19, 258-260.

Ayres, M.D., Howard, S. C., Kuzio, J., Ferber, M. L., and Possee, R. D. (1994). The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* 202, 586-605.

Becker, D., and Knebel-Morsdorf, D. (1993). Sequence and temporal appearance of the early tanscribed baculovirus gene HE65. *Journal of Virology* 67, 5867-5872.

Bilimoria, S. L. (1983). Genomic divergence among single-nucleocapsid nuclear polyhedrosis viruses of plusiine hosts. *Virology* 127, 15-23.

Black, B. C., Brennan, L. A., Dierks, P. M., Gard, I. E. (1997). Commercialization of baculoviral insecticides. In " *The Baculovirus*" (Miller L. K., ed). Plenum Press. New York and London. pp 341-387.

Blissard, G. W., Black, B., Crook, N., Keddie, B. A., Possee, R., Rohrmann, G., Theilmann, D. A., and Volkman, L. E. (2000). Baculoviridas. In "*Virus Taxonomy*" Seventh Report of International Committee on Taxonomy of Viruses" (Regenmortel, M. H. V., C. M. Fauquet, D. H., Bishop, E. B. L., Carstens, M. K., Estes, S. M., Lemon, J.,

https://etd.uwc.ac.za/

Maniloff, M. A., Mayo, D. J., McGeoch, D. J., Prongle, C. R., and Wickner, R. D. eds.). pp 195-202. Academic press, New York.

Bideshi, D. K., Bigot, Y., and Fedcrici, B. A. (2000). Molecular characterization and phylogenetic analysis of the *Harrisina brillians* granulovirus granulin gene. *Archive Virology* 145, 1933-1945.

Blissard, G. W., and Rohrmann, G. F. (1989). Location, sequence, transcriptional mapping and temporal expression of the gp64 envelope glycooprotein gene of the *Orgyia psendotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 170, 537-555.

Blissard, G. W., and G. F. Rohrmann. (1990). Baculovirus diversity and molecular biology. *Annual Review of Entomology* 35, 127-155.

Blissard, G. W., and Rohrmann, G. F. (1991). Baculovirus gp64 gene expression: Analysis of sequences modulating early transcription and transactivation by IE1. *Journal of Virology* 65, 5820-5827.

Blissard, G. W., P. H. Kogan, R. Wei, and G. F. Rohrmann. (1992). A synthetic early promoter from a baculovirus: roles of the TATA box and conserved start site CAGT sequence in basal level of transcription. *Virology* 190, 783-793.

Blissard, G. W. (1996). Baculovirus-insect cell interactions. *Cytotechnology* 20, 73-93.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Braunagel, S. C., and Summers, M. D. (1994). *Autographa californica* nuclear polyhedrsis virus, PDV and ECV viral envelopes and nucleocapsids: structural proteins, antigens, lipids and fatty acid profiles. *Virology* 202, 315-328.

Brent, R., snd Ptashne, M. (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell.* 43, 729-736.

Bulach, D. M., Kumar, C. A., Zaia, A., Liang, B., and Tribe, D. E. (1999). Group II nucleopolyhedrovirus subgroups revealed by phylogenetic analysis of polyhedrin and DNA polymerase gene sequences. *Journal of Invertebrate Pathology* 73, 59-73.

Carson, D. D., Guarino., L. A., and Summers, M. D. (1988). Functional mapping of an AcNPV immediate early gene which augments expression of the tans-activated 39K gene. *Virology* 162, 444-451.

Carson, D. D., Summers, M. D., and Guarino, L. A. (1991). Molecular analysis of a baculovirus regulatory gene. *Virology* 182, 279-286.

Carstens, E. B., Lu, A. L. and Chan, H. L. B. (1993). Sequence, transcriptional mapping, and overexpression of p47, a baculovirus gene regulation late gene expression. *Journal of Virology* 67, 2513-2520.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Chen, X., Ijkel, W. F. J., Dominy, C., Zanotto, P., Hashimoto, Y., Faktor, O., Hayakawa, T., Wang, C., Prekumar, A., Mathavan, S., Krell, P. J., Hu, Z., Vlak, J. M. (1999). *Virus Research* 65, 21-32.

Chen, X., Li, M., Sun, X., Hu., and Vlak, J. M. (2000). Genomic organization of *Helicoverpa armigera* single-nuclecapsid nucelopolyhedrovirus. *Archives of Virology* 145, 2539-2555

Chen, X., Sun, X., Hu, Z., Li, M., O'Reilly, D. R., Zuidema, D., and Vlak, J. M. (2000). Genetic engineering of Helicoverpa armigera single-nucleocapsid nucleohedrovirus as an improved pesticide. *Journal of Invertebrate Pathology* 76, 140-146.

Chen, X., Ijkel, W. F. J., Tarchini, R., Sun, X., Sandbrink, H., Wang, H., Peters, S., Zuidema, d., Lankhorst, R. K., Vlak, J. M., and Hu, Z. (2001). The sequence of the *Helicoverpa armigera* single nuclecapsid nucelopolyhedrovirus genome. *Journal of Genenal Virology* 82, 241-257.

Chen, X. (2001). The genomic and genetic engineering of *Helicoverpa armigera* nucleopolyhedroviruses. PhD thesis, Wageningen University.

Chen, X., Zhang, W. J., Wong, J., Chun, G., Lu, A., McCutchen, B. F., Presnail, J. K., Herrmann, R., Dolan, M., Tingey, S., Hu, Z. H., and Vlak, J. M. (2002). Comparative analysis of the complete genome sequences of *Helicoverpa zea* and *Helicoverpa*

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

argmigera single-nucleocapsid nucleopokyhedroviruses. *Journal of Genenal Virology* 83, 673-684.

Cheng, L. Z., Workman, J. L., Kingston, R. E., and Kelly, T. J. (1992). Regulation of DNA replication in vitro by the transcriptional activation domains of GAL4-VP16. *Proceeding of the National Academy of Sciences of USA* 89, 589-593.

Cherbas, L., and Cherbas, P. (1993). The arthropod initiator: the capsite consensus plays an important role in transcription. *Insect Biochemistry and Molocular Biology* 23, 81-90.

Chisholm, G. E., and Henner, D. J. (1988). Multiple early transcripts and spliceing of the *Autographa californica* nuclear polyhedrosis virus ie-1 gene. *Journal of Virology* 62, 3193-3200.

Clarke, E. E., Tristem, M., Cory, J. S., and O'Reilly, D. R. (1996). Characterization of the ecdysteroid UDP-glucosyltransferase gene from *Mamestra brassicae* nucleopolyhedrosis virus. *Journal of General Virology* 77, 2865-2871.

Cochran, M. A., and Faulkner, P. (1983). Location of homologous DNA sequences interspersed at five regions in the baculovirus AcMNPV genome. *Journal of Virology* 45, 961-970.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Condreay, J. P., Witherspoon, S. M., Clay, W. C., and Kost, T. A. (1999). Transient and stable gene expression in mammalian cells transduced with a recombinant baculovirus vector. *Proceeding of the National Academy of Sciences of USA* 96, 127-132.

Cory, J. S., Hirst, M. L., Williams, T., Hails, R. S., Goulson, D., and Green, B. M. (1994). Field trial of a genetically improved baculovirus insecticide. *Nature* 370, 138-140.

Cowan, P., Bulach, D., Goodge, K., Robertson, A., and Tribe, D. E. (1994). Nucleotide sequence of the polyhedrin gene region of Helicoverpa zea single nucleocapsid nuclear polyhedrosis virus: placement of the virus in lepidopteran nuclear polyhedrosis virus group II. *Journal of General Virology* 75, 3211-3218.

Crawford, A. M., and Miller, L. K. (1988). Characterization of an early gene accelerating expression of late genes of the baculovirus *Autographa californica* nuclear polyhedrosis virus. Journal of Virology 62, 2773-2781.

Croizier, L., Taha, A., Croizier, G., and Lopez Ferber, M. The complete sequence of the potato tuber moth, *Phthorimaea operculella* granulovirus. Unpublished.

Cui X., De Vivo I., Slany R., Miyamoto A., Firestein R., and Cleary M.L. (1998). Association of SET doamin and myotubularin-related proteins modulates growth control. *Nature Genetics* 18, 331-337.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Davis, T. R., and Wood, H. A. (1996). In vitro characterization of a *Trichoplusia ni* single nuclear polyhedrosis viruse. *Journal of General Virology* 77, 2303-2310.

Debolt, J. W., Wolf, W. W., Henneberry, T. J., and P. V. Vail. (1979). Evaluation of light traps and sex pheromone for control of cabbage looper and other lepidopterous insect pests of lettuce. *USDA Technical Bulletin* 1606.

Dickson, J. A., and Friesen, P. D. (1991). Identification of upstream promoter elements medicating early transcription from the 35,000-molecular-weight protein gene of *Autographa californica* nuclear polyhedrosis virus. *Journal of Virology* 65, 4006-4016.

Dornan, M. H., Sears, M. K., and Sears, J. G. (1995). Evaluation of a binomial model for insecticide application to control lepidopterous pests in cabbage. *Journal of Economical Entomology* 88, 302-306.

Efron B. CMBMS-NSF Regional Conference Series in Applied Mathematics Monograph 38, Society of Industrial and Applied Mathematics, Philadephia.

Engelhard, E. K., Kam-Morgan, L. N. W., Washburn, J. O., and Volkman, L. E. (1994). The insect tracheal system: A conduit for the systemic spread of *Autographa californica* M nuclear polyhedrosis virus. *Proceeding of the National Academy of Sciences of USA* 91, 3224-3227.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Erlandson, M., Li, S., Gillott, C., and Moody, D. (2002). Sequence and organization of the *Mamestra configurata* nucleopolyhedrovirus genome. *Virology* 294, 106-121.

Evans, H. F. (1981). Quantitative assessment of the relationship between dosage and response of the nuclear polyhedrosis virus of *Mamestra brassicae*. *Journal of Invertebrate Pathology* 37, 101-109.

Fielding, B. C., and Davison, S. (1999). The characterization and phylogenetic relationship of the *Trichoplusia ni* single capsid nuclear polyhedrosis viruse polyhedrin gene. *Virus Genes* 19, 67-73.

Fielding, B. C. (2001). The molecular characterization of a baculovirus isolated from *Trichoplusia ni*. PhD thesis, University of the Western Cape.

Fielding, B. C., Sehaam Khan., Wang, W., Kruger, C., Abrahams, R., and Davison, S. (2002). The genetic organization of a 2966 basepair DNA fragment and phylogeny of a single capsid nuclear polyhedrosis viruse isolated from *Trichoplusia ni*. *Virus Genes* 25, 35-43.

Freement P. S. (1993). The RING finger. A novel protein sequence motif related to the zinc finger. *Annals of the New York Academy of Sciences* 684, 174-192.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Friesen, P. D. (1997). Regulation of baculovirus early gene expression, In "*The Baculovviruses*". (Miller, L. K. ed), Plenum Press, New York/London, pp 141-170.

Fuchs, L. Y., Woods, M. S., and Weaver, R. F. (1983). Viral transcription during *Autographa californica* nuclear polyhedrosis virus infection: a novel RNA polymerase induced in infected *Spodoptera frugiperda* cells. *Journal of Virology* 43, 641-646.

Gard, I. E. (1997). Field testinga genetically modified virus. In " Microbial Insecticides: Novelty or Necessity?", BCPC symposium No.68, pp 101-114.

Garrity, D. B., Chang, M. J. and Blissard, G. W. (1997). Late promoter selection in the baculovirus *gp64* envelope fusion protein gene. *Virology* 231, 167-181.

Gershburg, E., Stockhom, D., Frog, O., Rashi, S., Gurevitz, M.,and Chejanovsky, N. (1998). Baculovirus-mediated expression of a scorpion depressant toxin improves the insecticidal efficacy achieved with excitatory toxin. *FEBS Letters* 422, 132-136.

Gibson, T. J., Aasland R., and Stewart, A. F. (1995). The PHD finger: implications for chromatin-emdiated transcriptional regulation. *Trends in Biochemical Sciences* 20, 56-59.

Gomi, S., Majima, K., and Maeda, S. (1999). Sequence analysis of the genome of *Bombyx mori* nucleopolyhedrovirus. *Journal of Genenal Virology* 80, 1323-1337.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Granados, R. R., and Lawler, K. A. (1981). *In vitro* pathway of *Autographa californica* baculovirus invasion and infection. *Virology* 90, 170-174.

Granados, R. R., Derksen, A. C. G., and Dwyer, K. G. (1986). Replication of the *Trichoplusia ni* granulosis and nuclear polyhedrosis viruses in cell cultures. *Virology* 152, 472-476.

Granados, R. R., Guoxun, L., Derksen, A. C. G., and Mckenna, K. A. (1994). A new insect cell line from *Trichoplusia ni* (BTI-Tn-5B1-4) susceptible to *Trichoplusia ni* single enveloped nuclear polyhedrosis virus. *Journal of Invertebrate of Pathology* 64, 260-266.

Green, S., Kumer, V., Theulaz, I. Wahli, W., and Chambon, P. (1988). The N-terminal DNA-binding 'zinc finger' of the oestrogen and glucocorticoid receptors determines target gene specificity. *EMBO J.* 7, 3037-3044.

Guarino, L. A., Gonzalez, M. A., and Summers, M. D. (1986a). Complete sequence and enhancer function of the homologous DNA regions of *Autographa californica* nuclear polyhedrosis virus. *Journal of Virology* 60, 224-229.

Guarino, L. A., and Summers, M. D. (1986b). Interspersed homologous DNA of *Autographa californica* nuclear polyhedrosis virus enhances delayed-early gene expression. *Journal of Virology* 60, 215-223.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Guarino, L. A., and Summers, M. D. (1986c). Functional mapping of a trans-activiting gene required for expression of a baculovirus delayed-early gene. *Journal of Virology* 57, 563-571.

Guarino, L. A., and Smith, M. W. (1990). Nucleotide sequence of the 39K gene region of *Autographa californica* nuclear polyhedrosis virus. *Virology* 179, 1-8.

Guarino, A. L., and Dong, W. (1991). Expression of an enhancer-binding protein in insect cells transfected with the *Autographa californica* nuclear polyhedrosis virus IE1 gene. *Journal of Virology* 65:3676-3680.

Guarino, L. A., and Smith, M. (1992). Regulation of delayed-early gene transcription by dual TATA boxes. *Journal of Virology* 66, 3733-3739.

Guarino, A.L., and Dong, W. (1994). Functional dissection of the *Autographa californica* nuclear polyhedrosis virus enhancer element hr5. *Virology* 200, 328-335.

Guarino, A.L., Xu, B., Jin, J. P., and Dong, W. (1998). A virus-encoded RNA polymerase purified from baculovirus-infected cells. *Journal of Virology* 72, 7985-7991.

Guilfoyle, R. A., Osheroff, W. P., and Rossini, M. (1985). Two functions encoded by adenovirus early region 1A are responsible for the activation and repression of the DNA-binding protein gene. *EMBO Journal*. 4, 707-713.

https://etd.uwc.ac.za/

Hammock, B. D., Bonning, B. C., Possee, R. D., Hanzlik, T. N., Maeda, S. (1990). Expression and effects of juvenile hormone esterase in a baculovirus vector. *Nature* 344, 458-461.

Hannenhalli, S., Chappey, C., Koonin, E. V., and Pevzner, P. A. (1995). Genome sequence comparison and scenarios for gene rearrangements-a test-case. *Genomics*, 30, 299-311.

Hashimoto, Y., Hayakawa, T., Ueno, Y., Fujita, T., Sano, Y., and Matsumoto, T. (2000). Sequence analysis of the *Plutella xylostella* granulovirus genome. *Virology* 275, 358-372.

Hawtin, R. E., Zarkowska, T., Arnold, K. A., Thomas, C. A., Gooday, G. W., Kin, L. A., Possee, R. D. (1997). Liquefaction of the *Autographa californica* nuclear polyhedrovirus-infected insect is dependent of the integrity of virus-encoded chitinase and cathepsin genes. *Virology* 238, 243-253.

Hayakawa, T., Ko, R., Okano, K., Seong, S., Goto, C., and Maeda, S. (1999). Sequence analysis of the *Xestia c-nigrum* granulovirus genome. *Virology* 262, 277-297.

Hayakawa, T., Rohrmann, G., and Hashimoto, Y. (2000). Pattern of genome organization and content in lepidoptern baculoviruses. Virology 278, 1-12.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

He, Z., Brinton, B. T., Greenblatt, J., Hassell, J. A., and Ingles, C. J. (1993). The transactivator proteins VP16 and GAL4 bind replication factor A. *Cell* 73, 1223-1232.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28, 351-359.

Hernandez-Crespo, P., Hails, R. S., Sait, S. M., Green, B. M., Carty, T. M., and Cory, J. S. (1999). Response of hosts of varying susceptibility to a recombinant baculovirus insecticide in the field. *Biological Control* 16, 119-127.

Herniou, E. A., Luque, T., Chen, X., Vlak, J. M., Winstanley, D., Cory, J. S., and O'Reilly D. (2001). Use of whole genome sequences data to infer baculovirus phylogeny. *Journal of Virology* 75, 8117-8126.

Herniou, E. A., Olszewski, J. A., Cory, J. S., and O'Reilly, D. R. (2003). The genome sequence and evolution of baculovirus. *Annual Review Entomology* 48, 211-234.

Hill, J. E., Kuzio, J., Faulkner, P. (1995). Identification and characterization of the v-cath gene of the baculovirus CfMNPV. *Biochemical and Biophysical Acta*. 1264, 275-278.

Hoffman, M. P., and Fordsham, A. C. (1993). Nature enemies of vegetable insect pests. New York, Cornell Cooperative Extension Publication. pp 5.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Hofmann, C., Sandig, V., Jennings, G., Roudolph, M., Schlag, P., and Strauss, M. (1995) Efficient gene transfer into human hepatocytes by baculovirus vectors. *Proceedings of the National academy of Science of USA* 92, 10099-10103.

Hope, I. A., and Struhl, K. (1986). Functional dissectin of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell 46, 885-894.

Hoopes, R. R., and Rohrmann, G. F. (1991). In vitro transcription of baculovirus immediate early genes: accurate mRNA initiation by nuclear extracts from both insect and human cells. *Proceedings of the National academy of Science of USA* 88:4513-4517.

Horton, H. M., and Burand, J. P. (1993). Saturable attachment sites for polyhedron-derived baculovirus on insect cells and evidence for entry via direct membrane fusion. *Journal of Virology* 67, 1860-1868.

Hu, Z. H., Arif, B. M., Jin, F., Martens, J. W. M., Chen, X. W., Sun, J. S., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1998). Distinct gene arrangement in the *Buzura suppressaria* single nucleocapsid nucleopolyhedrovirus genome. *Journal of General Virology* 79, 2841-2851.

Hu, Y. F., Hao, Z. L., and Li, R. (1999). Chromatin remodeling and activation of chromosomal DNA replication by an acidic transcriptional activation domain from BRCA1. *Genes & Development* 13, 637-642.

https://etd.uwc.ac.za/

Hyink, O., Dellow, R. A., Olsen, M., Cradoc-Davies, K. M. B., Drake, K., Herniou, E. A., Cory, J. S., O'Reilly, D. R., and Ward, V. K. (2002). Whole genome analysis of the *Epiphyas postvittana* nucleopolyhedrovirus. *Journal of Genenal Virology* 83, 957-971.

Ijkel, W. F. J. (2001). The genome of *Spodoptera exigua* multicapsid nucleopolyhedrovirus: a study on unique features. PhD thesis, Wageningen University.

IJKel, W. F. J., van Strien, E. A., Heldens, J. G. M., Broer, R., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1999). Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome. *Journal of Genenal Virology* 80, 3289-3304.

Jarrell, K. A., and Meselson, M. (1991). Drosophila retrotransposon promoter includes an essential sequence at the initiation site and requires a downstream sequence for full activity. *Proceedings of the National academy of Science of USA* 88, 102-104.

Jarvis, D. L, and Summers, M. D. (1989). Glycosylation and secretion of human tissue plasminigen activator in recombinant baculovirus-infected insect cells. *Molecular and Cellular biology* 9, 214-223.

Jarvis, D. L., Fleming, J. G. W., Kovacs, G. R., Summers, M. D., and Guarino, L. A. (1990). Use of early baculovirus promoters for continuous expression and efficient

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

processing of foreign gene products in stable-transformed lepidopteran cells. *Bio/Thchnology* 8, 950-955.

Jarvis, D. L., Reilly, L. M., Hoover, k., Schultz, C., and Hammock, B. D. (1996a). Construction and characterization of immediate early baculovirus pesticides. *Biological Control* 7, 228-235.

Jarvis, D. L., Weinkauf, C., and Guarino, L. A. (1996b). Immediate-early baculovirus vectors for foreign gene expression in transformed or infected insect cells. *Protein Expression and Purification* 8, 191-203.

Jenuwein T., Laible G., Dorn R., and Reuter G. (1998). SET domain proteins modulate chroamtin domains in eu- and heterochromatin. *Cellular and Molecular Life Sciences* 54, 80-93.

Jin, T., Qi, Y., Liu, D., and Su, F. (1999). Nucleotide sequence of a 5892 base pairs fragment of the LsMNPV genome and phylogenetic analysis of LsMNPV. Virus Genes 18, 265-276.

Kang, W., Tristern, M., Maeda, S., Crook, N. E., and O'Reilly, D. R. (1998). Identification and characterization of the *Cydia pomonella* granulovirus cathepsin and chitinase genes. *Journal of General Virology* 79, 2283-2292.

https://etd.uwc.ac.za/

Keddie, B. A., Aponte, G. W., and Volkman, L. E. (1989). The pathway of infection of *Autographa californica* nucleocapsid polyhedrosis virus in an inset host. *Science* 243, 1728-1730.

King, L. A., and Possee, R. D. (1990). A baculovirus expression vector derived from the basic protein promoter of *Autographa californica* nuclear polyhedrosis virus. *Journal of Genenal Virology* 71, 971-976.

Kishaba, A. N., Whitaker, T. W., Berry, W., and Toba, H. H. (1976). Cabbage looper oviposition and survival of progeny on leafy vegetables. *Hortscience* 11, (3). 216-217.

Kits, P. A., Kin, L. A., Possee, R. D. (1995). Identification and preliminary characterization of a chitinase gene in the *Autographa californica* nucleocapsid polyhedrosis virus. *Virology* 212, 673-685.

Knebel-Morsdorf, D., Kremer, A., and Jahnel, F. (1993). Baculovirus gene ME53, which contains a putative zinc finger motif, is one of the majar early-transcribed genes. *Journal of Virology* 67, 753-758.

Knebel-Morsdorf, D., Flipsen, J. T. M., Roncarati, R., Jahnel, F., Kleefsman, A. W. F., and Vlak, J. M. (1996). Baculovirus infection of Spodoptera exigua larvae: lacz expression driven by promoters of early genes pe38 and me53 in larvval tissue. *Journal of General Virology* 77, 815-824.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Kogan, P. H. and Blissard G. W. (1994). A baculovirus gp64 early promoter is activated by host transcription factor binding to CACGTG and GATA elements. *Journal of Virology* 68, 813-822.

Kogan, P. H., Chen, X., and Blissard, G. W. (1995). Overlapping TATA-dependent and TATA-independent early promoter actvities in the baculovirus gp64 envelope fusion protein gene. *Journal of Virology* 69, 1452-1461.

Kool, M., Van Den Berg, P. M. M. M., Tramper, J., Goldbach, R. W.,and Vlak, J. M. (1993a). Location of two putative origins of DNA replication of *Autographa californica* nuclear polyhedrosis virus. *Virology* 192, 94-101.

Kool, M., Voeten, J. T. M., Goldbach, R. W., Tramper, J., and Vlak, J. M. (1993b). Identification of seven putative origins of *Autographa californica* multiple nuclear polyhedrosis virus DNA replication. *Journal of General Virology* 74, 2661-2668.

Kool, M., Ahrens, C. H., Goldbach, R. W., and Rohrmann, G. F. (1994). Identification of genes involved in DNA replicatin of *Autographa californica* baculovirus. *Proceeding of the National Academy of Sciences of USA* 91, 11212-11216.

Kool, M., Ahrens, C.H., Vlak. J. M., and Rohrmann, G.F. (1995). Replication of baculovirus DNA. *Journal of General Virology* 76, 2103-2118.

UNIVERSITY of the
WESTERN CAPE
https://etd.uwc.ac.za/

Korst, T. A., and Condreay, J. P. (1999). Recombinant baculovirus as expression vectors for insect and mammalian cells. *Current Opinion in Biotechnology* 10, 428-433.

Kovacs, G. R., Guarino, L. A., and Summers, M. D. (1991). Novel regulatory properties of the IE1 and IE0 transactivators encoded by the baculovirus *Autographa californica* multicapsid nuclear polyhedrosis virus. *Journal of Virology* 65, 5281-5288.

Kovacs G. R., Choi, J., Guarino, L. A., and Summers, M. D. (1992). Functional dissection of the *Autographa californica* nuclear polyhedrosis virus immediate-early 1 transcriptional regulatory protein. *Journal of Virology* 66, 7429-7439.

Krappa, R., and Knebel-Morsdorf, D. (1991). Identification of the very early transcribed baculovirus gene PE-38. *Journal of Virology* 65, 805-812.

Kremer, A., and Knebel-Morsdorf, D. (1998). The early baculovirus he65 promoter: on the mechanism of transcriptional activation by IE1. *Virology* 249, 336-351.

Kumar, V., and Chambon, P. (1988). The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. *Cell* 55, 145-156.

https://etd.uwc.ac.za/

Kuzio, J., Pearson, M. N., Harwood, S. H., Funk, C. J., Evans, J. T., Slavicek, J. M., and Rohrmann, G. F. (1999). Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria dispar. *Virology* 253, 17-34.

Leisy, D. J., and Rohrmann, G. F. (1993) Characterization of the replication of plasmids containing hr sequences in baculovirus-infected *Spodoptera frugiperda* cells. *Virology* 196, 722-730.

Leisy, D. J., Rasmussen, C., Kim, H. T. and Rohrmann, G. F. (1995). The *Autographa californica* nuclear polyhedrosis virus homologous region 1a: identical sequences are essential for DNA replication activity and transcriptional enhancer function. *Virology*. 208, 742-752.

Leisy, D. J., C. Rasmussen, E. O. Owusu, and Rohrmann, G. F. (1997). A mechanism for negative gene regulation in *Autographa californica* multi-nucleocapsid nuclear polyhedrosis virus. *Journal of Virology* 71, 5088-5094.

Li, Y., Passarelli, A. L., and Miller, L. K. (1993). Identification, sequences, and transcriptional mapping of *lef-3*, a baculovirus gene involved in late and very late gene expression. *Journal of Virology* 67, 5260-5268.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Li, R., Yu, D. S., Tanaka, M., Zheng, L., Berger, S. L., and Stillman B. (1998). Activation of chromosomal DNA replication in Saccharomyces cerevisiae by acidic trancriptional activation domains. *Molecular Cell Biology* 18, 1296-1302.

Li, C. B., Li, Z. F., Yan, Q. S., Ilu, G. D., and Pang, Y. (2001). Cloning and sequencing of *Spodoptera litura* multicapsid nucleopolyhedrovirus gp37 gene cluster. *Progress in Biochemistry & Biophysics* 28, 677-682.

Lin, G., and Blissard, G. W. (2002). Analysis of an *Autographa californica* nucleopolyhedrovirus *lef-11* knockout: LEF-11 is essential for viral DNA replication. *Journal Virology* 76, 2770-2779.

Liu, J., and Maruniak, J. E. (1999). Molecular characterization of genes in the GP41 region of baculovirus and phyologenetic analysis based on GP41 and polyhedrin genes. *Virus Research* 64, 187-196.

Locke, M. Nucleolar cycles during the fifth stadium in Manduca epidermis. (1985). *Tissue Cell* 17, 901-921.

Lu, A., and Carstens, E. B. (1991). Nucleotide sequence of a gene essential for viral DNA replication in the baculovirus *Autographa californica* nuclear polyhedrosis virus. *Virology* 181, 336-347.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Lu. A., and Carstens, E. B. (1993). immediate-early baculovirus genes transactivate the p143 gene promoter of *Autographa californica* nuclear polyhedrosis virus. *Virology* 195, 710-718.

Lu, A., and L. K. Miller. 1995. The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication. *Journal of Virology* 69, 975-982.

Lu, A., Krell, P. J., Vlak, J. M., and Rohrmann, G. F. (1997). Baculovirus DNA replication. In "*The Baculovviruses*". (Miller, L. K. ed), Plenum Press, New York, pp 171-186.

Lu, A., and Miller, L. K. (1997). Regulation of baculovirus late and very late gene expression, In "*The Baculovviruses*". (Miller, L. K. ed), Plenum Press, New York, pp 193-211.

Luckow, V. A. (1991). Cloning and expression of heterologous genes in insect cells with baculovirus vector. in "*Recombinant DNA Technology and Application* ". (Prokop, A., Bjapai, R. K. and Ho, C.S. ed). McGraw-Hill, New York, . pp 97-152.

Luckow, V. A., Lee, S. C., Barry, G. F. and Olins, P.O. (1993). Efficient generation of infectous recombinant baculoviruses by site-specific transposon-medicated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli. *Journal of Virology* 67, 4566-4579.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

Luque, T., Finch, R., Crook, N., O'Reilly, D. R., and Winstanley, D. (2001). The complete sequence of the *Cydia pomonella* granulovirus genome. *Journal of Genenal Virology* 82, 2531-2547.

Maeda, S., Volrath, S. L., Hanzlik, T. N., Harper, S. A., Majima, K., Maddox, D. W., Hammock, B. D.,and Fowler, E. (1991). Insecticidal effects of an insect-specific neurotoxin expressed by a recombinant baculovirus. *Virology* 184, 777-780.

Mans, R. M. W., and Knebel-Morsdorf, D. (1998). In vitro trancription of pe38/polyhedrin hybrid promoters reveals sequences essential for recognition by the baculovirus-induced RNA polymerase and for the strength of very late viral promoters. *Journal of Virology* 72, 2991-2998.

Margot P., Bjornson, R., Pearson, G., and Rohrmann, G. (1992). The *Autographa californica* baculovirus genome: Evidence for multiple replication origins. *Science* 257, 1382-1384.

Marsden, D. A. (1979). Insect Pest Series No. 6 Cabbage Worm. Cooperative Extension Service, University of Hawaii, College of Tropical Agriculture and Human Resources.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Martens, J. W. M., Van Oers, M. M., Van de Bilt, B. D., Oudshoorn, P., and Vlak, J. M. (1995). Development of a baculovirus vector that facilitates the generation of o10-based recombinant. *Journal of Virological Methods* 52, 15-19.

Massari, M. E., and Murre, C. (2000). Helix-loop-helix proteins: Regulators of transcription in eukaryotic organisms. *Molecular Cell Biology* 20:429-440.

McCutchen, B. F., Choudary, P. V., Crenshaw, R., Maddox, D. W., Kamita, S. G., Palekar, N., Volrath, S., Fowler, E., Hammock, B. D., and Maeda, S. (1991). Development of recombinant baculovirus expressing an insect selective neurotoxin potential for pest control. *Biotechnology* 9, 848-852.

McDougal, V. V., and Guarino, L. A. (2000). The *Autographa californica* nuclear polyhedrosis virus p143 gene encodes a DNA helicase. *Journal of Virology* 74, 5273-5279.

McLachlin, J. r., and Miller, L. K. (1994). Identification and characterization of vlf-1, a baculovirus gene involved in very late gene expression. *Journal of Virology* 68, 7746-7756.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Miller, L.K., and Dawes, K.P., (1978). Restriction endonuclease analysis for the identification of baculovirus pesticides. *Applied and Environmental Microbiology* 35, 411-421.

Miller, J., McLachlan, A. D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. *EMBO Journal* 4, 1609-1614.

Mitchell, P. J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245, 371-378

Monsma, S. A., Oommens, A. G. P., and Blissard, G. W. (1996). The gp64 envelope fusion protein is an essential baculovirus protein required for cell to cell transmission of infection. *Journal of Virology* 70, 4607-4616.

Montague, M. G., and Hutchison, C. A. (2000). Gene content phylogeny of herpesviruses. *Proceeding of the National Academy of Sciences of USA* 97, 5334-5339.

Morris, T. D., and Miller, L. K. (1994). Mutational analysis of a baculovirus major late promoter. *Gene* 140, 147-153.

Morris, T. D., Todd, J. W., Fisher, B., and Miller, L. K. (1994). Identification of *lef-7*, a baculovirus gene affecting late gene expression. *Virology* 200, 360-369.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Moser, B. A., Becnel, J. J., White, S. E., Afonso, C., Kutish, G., Shanker, S., and Almira, E. (2001). Morphological and molecular evidence that *Culex nigripalpus* baculovirus is an unusual member of the family Baculoviridae. *Journal of General Virology* 82, 283-297.

Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers M. D. (1995). Virus Taxonomy. *Sixth Report of the International Committee on Taxonomy of Viruses*. Springer-Verlag, Wein/New York, pp 104-113.

Murphy, C. I., Lennick, M., Lehar, S. M., Beltz, G., A., and Young, E. (1990). Temporal expression of HIV-1 envelope proteins in baculovirus-infected insect cells: Implication for glycosylation and CD4 binding. *Genetic Analysis Techniques and Applications* 7, 160-171.

Murre, C., McCaw, P. S., and Baaltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless, MyoD*, and *myc* Proteins. *Cell.* 56, 777-783.

Nissen, M. S., and Friesen, P. D. (1989). Molecular analysis of the transcriptional regulatory region of an early baculovirus gene. *Journal of Virology* 63, 493-503.

https://etd.uwc.ac.za/

Novina, C. D., and Roy, A. L. (1996). Core promoters and transcriptional control. *Trends in Genetics* 12, 351-355.

Ohkawa, T., Majima, K., and Maeda, S. (1994). A cystein protease encoded by the baculovirus *Bombyx mori* nuclear polyhedrosis virus. *Journal of Virology* 68, 6619-6625.

Ohresser, M., Morin, N., Cerutti, M., and Delsert, C. (1994). Temporal regulation of a complex and unconventional promoter by viral products. *Journal of Virology* 68, 2589-2597.

Oomens, A. G. P., and Blissard, G. W. (1999). Requirement for Gp64 to drive efficient budding of *Autographa californica* multicapsid nucleopolyhedrovirus. *Virology* 254, 297-314.

Oomens A. G. P., Monsma, S. a., and Blissard, G. W. (1995). The baculovirus GP64 envelope fusion protein: synthesis, oligomerization, and processing. *Virology* 209, 592-603.

Ooi, B. G., Rankin, C., and Miller, L. K. (1989). Downstream sequences augment transcription from the essential initiation site of a baculovirus polyhedrin gene. *Journal of Molecular Biology* 210, 721-736.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/

O'Reilly , D. R., and Miller, L. K. (1991) Improvement of a baculovirus pesticide by deletion of the *egt* gene. *Bio/Thchnology* 9, 1086-1089.

O'Reilly, D. R., Miller, L. K., and Luckow, V. A. (1992). Baculovirus expression vector: a laboratory manual. W. H. Freeman and Co., New York.

Olson, V. A., Wetter, J. A., and Friesen, P. D. (2001). Oligomerization mediated by a helix-loop-helix-like Domain of baculovirus IE1 is required for early promoter transactivation. *Journal of Virology* 75, 6042-6051.

Pang, Y., Yu, J., Wang, L., Hu, X., Bao, W., Li, G., Chen, C., Han, H., Hu, S., and Yang, H. (2001) Sequence analysis of the *Spodopter litura* multicapsid nucleopolyhedrovirus genome. *Virology* 287, 391-404.

Passarelli, A. L., and Miller, L. K. (1993a). Three baculovirus genes involved in late and very late gene expression: ie-1, ie-n, and lef-2. *Journal of Virology* 67, 2149-2158.

Passarelli, A. L., and Miller, L. K. (1993b). Identification and characterization of *lef-1*, a baculovirus gene involved in late and very late gene expression. *Journal of Virology* 67, 3481-3488.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Passarelli, A. L., and Miller, L. K. (1993c). Identification and of genes encoding late expression factors located between 56.0 and 64.4 map units of the *Autographa californica* nucleocapsid nuclear polyhedrosis virus genome. *Virology* 197, 704-714

Passarelli, A. L., and Miller, L. K. (1994). Identification and transcriptional regulation of the baculovirus *lef-6* gene. *Journal of Virology* 68, 4458-4467.

Passarelli, A. L., Todd, J. W., and Miller, L. K. (1994). A baculovirus gene involved in late gene expression predicts a large polypeptide with a conserved motif of RNA polymerases. *Journal of Virology* 68, 4673-4678.

Pathakamuri, J. A., and Theilmann, D. A. (2002). The acidic activation domain of the baculovirus transactivator IE-1 contains a virus-specific domain essential for DNA replication. *Journal of Virology* 76, 5598-5604.

Patikoglou, G., and S. K. Burley. (1997). Eukaryotic transcription factor-DNA complexes. *Annual Review of Biophysics and Biomolecular Structure* 26, 289-325.

Pearson, M., Bjornson, R., Pearson, G., and Rohrmann, G. F. (1992). *Autographa californica* baculovirus genome: evidence for multiple replication origins. *Science* 257, 1382-1384.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Pearson, M. N., R. M. Bjornson, C. Ahrens, and Rohrmann, G. F. (1993). Identification and chracterization of a puptative origin of DNA replication in the genome of a baculovirus pathogenic for *Orgyia pseudotsugata*. *Vriology* 197:715-725.

Pearson, M., and Rohrmann, G. F. (1997). Splicing is required for transactivation by the immediate early gene 1 of the *Lymantria dispar* multinucleocapsid nuclear polyhedrosis virus. *Virology* 235, 153-165.

Persley, G. J. (ed). (1996). Biotechnology and Integrated Pest Management. CAB International, Oxon, U.K., pp 475.

Possee, R. D. (1997). Baculovirus as expression vectors. *Current Opinions in Biotechnology*. 8, 569-572.

Prikhod'ko E. A., and Miller, L. K. (1996). Induction of apoptosis by baculovirus transctivator IE-1. *Journal of Virology* 70, 7116-7124.

Prikhod'ko E. A., and Miller, L. K. (1999). The baculovirus PE38 protein augments apoptosis induced by transctivator IE-1. *Journal Virology* 73, 6691-6699.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. *Nature* (London) 335, 683-689.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Pullen, S. S., and P.D. Friesen. (1995a) Early transription of the IE-1 transregulator gene of *Autographa californica* nuclear polyhedrosis virus is regulated by DNA sequences within its 5' noncoding leader region. *Journal of Virology* 69:156-165.

Pullen, S. S., and Friesen, P. D. (1995b). The CAGT motif fonctions as an initiator element during early transcription of the baculovirus transregulator ie-1. *Journal of Virology* 69, 3575-3583.

Rankin, C., Ooi, B. G., and Miller, L. K. (1988). Eight base pairs encompassing the transcriptional start point are the major determinant for baculovirus polyhedrin gene expression. *Gene* 70, 39-49.

Rapp, J. C., Wilson, J. A., and Miller, L. K. (1998). Nineteen baculovirus open reading frame, including *lef-12*, support late gene expression. *Journal of Virology* 72, 10197-10206.

Ribeiro, B. M., Hutchinson, K., and Miller, L. K. (1994). A mutant baculovirus with a temperature-sensitive Ie-1transregulatory protein. *Journal of Virology* 68, 1078-1084.

Ricciardi, R. P., Jones, R. L., Cepko, C. L., Sharp, P. A., Roberts, B. E. (1981). Expression of early adenovirus genes requires a viral encoded acidic polypeptide. *Proceedings of the National academy of Science of USA* 78, 6121.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Robison, K.A., and Lopes, J. M. (2000). Survey and summary: Saccharomyces cerevisiae basic helix-loop-helix proteins regulate diverse biological processes. *Nucleic Acids Research* 28, 1499-1505.

Rodems, S. M., and P. D. Friesen. (1993) The hr5 transcriptional enhancer stimulates early expression from the *Autographa californica* nuclear polyhedrosis virus genome but is not required for virus replication. *Journal of Virology* 67, 5776-5785.

Rodems, S. M., and Friesen, P. D. (1995) Transcriptional enhancer activity of hr5 requires dual palindrome half sites that mediated binding of a dimeric form of the baculovirus transregulator IE1. *Journal of Virology* 69, 5368-5375.

Rodems S. M., Pullen, S. S., and Friesen, P. D. (1997). DNA-dependent transregulation by IE1of the *Autographa californica* nuclear polyhedrosis virus: Domains required for transactivation and DNA binding. *Journal of Virology* 71, 9270-9277.

Rohrmann, G. F. (1992) Baculovirus structure proteins. *Journal of Genenal Virology* 73, 749-761.

Saarinen, M. A., Troutner, K. A., Gladden, S. G.,Mitchell, C. M., and Murhammer, D. W. (1999). Recombinant protein synthesis in *Trichoplusia ni* BTI-Tn-5B1-4 insect cell aggregates. *Biotechnology Bioengineering* 63, 612-617.

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

Sadowski, I., Ma, J., Triezenberg, S., and Ptashne, M. (1988). Gal4-vP16 is an unusually potent transcriptional activator. *Nature* (London), 335, 563-564.

Sambrook, J., E. F. Fritsch., and T. Maniatis (1989). Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sambrook, J., and Russell, D. W. (2001) Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sarkis, C., Serguera, C., Peters, S., Buchet, d., Ridet, J. L., Edelman, L., Mallet, J. (2000). Effiicent transduction of neural cells in vitro and in vivo by a baculovirus-derived vector. *Proceedings of the National academy of Science of USA* 97, 14638-14643.

Slack, J. M., Kuzio, J., Faulkner, P. (1995). Characterization of v-cath, a cathepsin L-like proteinase expressed by baculovirus *Autographa californica* multiple nuclear polyhedrosis virus. *Journal of General Virology* 76, 1091-1098.

Slack, J. M., and Blissard, G. W. (1997). Identification of two independent transcritional activation domains in the *Autographa californica* multiple nucleocapsid nuclear polyhedrosis virus IE1 protein. *Journal of Virology* 71: 9579-9587.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Slavicek, J. M., Popham, H. J. R., Riegel, C. I. (1999). Deletion of the *Lymantria dispar* multicapsid nucleopolyhedrovirus ecdysteroid UDP-glucosyl transferase gene enhances viral killing speed in the last instar of the gypsy moth. *Biological Control* 16, 91-103.

Smith, G. E., and Summers, M. D. (1978). Analysis of baculovirus genome with restriction endonucleases. *Virology* 89: 517-27.

Smith, G. E., Summers, M. D., and Fraser, M. J. (1983). Production of human beta interferon in insect cells infected with baculovirus expression vector. *Molecular and Cellular biology* 3, 2156-2165.

Stewart, L. M., Hirst, M., Ferber, M. L., Merryweather, A. T., Cayley, P. J., and Possee, R. D. (1991). Construction of an improved baculovirus insecticide containing an insect-specific toxin gene. *Nature* 352, 85-88.

Swofford, D. L. 2000 PAUP Version 4, Sinauer Associates, Sunderland, Massachusetts.

Theilmann, D. A., and Stewart, S. (1991). Identification and characterization of the IE-1 gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 180, 492-508.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Theilmann, D. A., and Stewart, S. (1992). Molecular analysis of the transactivating IE-2 gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 187, 84-96.

Theilmann, D. A., Willis, L. G., Bosch, B., Forsythe, I. J., and Li, Q. (2001). The baculovirus transcriptional transactivator ie0 produces multiple products by internaaal initiation of translation. *Virology* 290, 211-223.

Thompson, J. d., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignement through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22, 4673-4680.

Todd, J. W., Passarelli, A. L., and Miller, L. K. (1995). Eighteen baculovirus genes, including *lef-11*, *p35, 39k*, and *p47*, support late gene expression. *Journal of Virology* 69, 968-974.

Todd, J. W., Passarelli, A. L., Lu, A., and Miller, L. K. (1996). Factors regulating baculovirus late and very late gene expression in transient-expression assays. *Journal of Virology* 70, 2307-2317

Tomarski, M. D.,and Miller, L. K. (1991). Insect paralysis of baculovirus-mediated expression of a mite neurotoxin gene. *Nature* 352, 82-85.

UNIVERSITY of the
WESTERN CAPE

https://etd.uwc.ac.za/

Tomarski, M. D.,and Miller, L. K. (1992). Expression of a paralysis neurotoxin to improve insect baculovirus as insecticides. *Boitechnology* 10, 545-549.

Triezenberg, S. J., Kingsbury, R. C., and Mcknight, S. L. (1988). Functional dessection of VP16, the trans-activator of herpes simplex virus immediate early gene expression, *Genes and Development* 2, 718-729.

Van Strien, E. A. (1997). Characterization of Spodoptera exigua baculovirus genome: structure and functional analysis of a 20 kb fragment. PhD thesis, Wageningen University.

Van Der Beek, C. P. (1980). On the origin of the polyhedral protein of the nuclear polyhedrosis virus of *Autographa californica*. PhD thesis. Agriculture University Wageningen.

Van Strien, E. A., Ijkel, W. F. J., Gerrits, H., Vlak J. M., and Zuidema, D. (2000). Characteristics of the transactivator gene *ie-1* of *Spodoptera exigua* multiple nucleopolyhedrovirus. *Archives of Virology* 145, 2115-2133.

Volkman, L. E., Goldsmith, P. A., Hess, R. T., and Faulkner, P. (1984). Nuetralization of budded *Autographa californica* NPV by a monoclonal antibody: Identification of the target antigen. *Virology* 133, 354-362.

UNIVERSITY of the WESTERN CAPE
https://etd.uwc.ac.za/

Wang, W., Leat. N., Fielding, B., and Davison, S. (2001). Identification, sequence analysis, and phylogeny of the immediate early gene 1 of the *Trichoplusia ni* single capsid nuclear polyhedrosis virus. *Virus Genes* 23, 53-62.

Weinberger, C., Hollenberg, S. M., Rosenfeld, G. M., and Evans R. M. (1985). Domain structure of human glucocorticoid receptro and its relationship to the v-erb-A oncogene product. *Nature* (London) 318, 670.

Wickham, T. J., Granados, R. R., Shuler, M. L., and Wood, H. A. (1992). Screening of insect cell lines for the productioon of recombinant proteins and infectious virus in the baculovirus expression system. *Biotechnology Progress* 8, 391-396.

Wickham, T. J., and Nemerow, G. R. (1993). Optimization of growth methods and recombinant protein production in BTI-Tn-5B1-4 insect cells using baculovirus expression system. *Biotechnology Progress.* 9, 25-30.

Xu, B., Yoo, S., and Guarino, L. A. (1995). Differential transcription of baculovirus late and very late promoters: fractionation of nuclear extracts by phosphocellulose chromatography. *Journal of Virology* 69, 2912-2917.

UNIVERSITY of the
WESTERN CAPE
https://etd.uwc.ac.za/

Zanotto, P. M. A., Kessing, B. D., and Maruniak, J. E. (1993). Phylogenetic interrelationships among baculoviruses: evolutionary rates and host associations. *Journal of Invertebrate Pathology* 62, 147-164.

https://etd.uwc.ac.za/

List of Publications

Wang, W., Leat. N., Fielding, B., and Davison, S. (2001). Identification, sequence analysis, and phylogeny of the immediate early gene 1 of the *Trichoplusia ni* single capsid nuclear polyhedrosis virus. *Virus Genes* 23, 53-62.

Fielding, B. C., Sehaam Khan., Wang, W., Kruger, C., Abrahams, R., and Davison, S. (2002). The genetic organization of a 2966 basepair DNA fragment and phylogeny of a single capsid nuclear polyhedrosis viruse isolated from *Trichoplusia ni*. *Virus Genes* 25, 35-43.

Wang, W., Davison, S., and Krell, P. (2003) Identification and characterization of a major early-transcribed gene of *Trichoplusia ni* single capsid nucleopolyhedrovirus using the baculovirus expression system. *Virus Research*. Submitted.

Wang, W., and Davison, S. Experimental Prediction of a DNA-Protein Binding Region of the *Trichoplusia ni* Single Nucleocapsid Nucleopolyhedrvirus IE-1. In preparation.

UNIVERSITY of the WESTERN CAPE

https://etd.uwc.ac.za/