UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Molecular cloning and expression of type C and D neurotoxin genes of *Clostridium botulinum*

by

Jennifer Rossouw

Submitted in partial fulfillment of the requirements for the degree of Magister Scientiae

In the Department of Microbiology and Plant Pathology
Faculty of Natural and Agricultural Sciences
University of Pretoria
Pretoria

JANUARY 2001

© University of Pretoria

DECLARATION

I declare that the dissertation, which I hereby submit for the degree Magister Scientiae at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at another university.

JRossouw
JENNIFER ROSSOUW

24/5/2001
DATE

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to the following people.

Prof L.H. Nel for his guidance and supervision during the course of this study and for the opportunity he granted me to continue my studies in molecular microbiology.

Dr. J. Theron for his guidance, support and motivation during the course of this study.

The Technology and Human Resources for Industry Programme (a partnership programme funded by the Department of Trade and Industry (DTI) and managed by the National Research Foundation (NRF)) and Design Biologix cc. for research grants.

Karen Nel for her interest, advice and help during this study. Pieter Smuts for anaerobic cultivation of *Clostridium botulinum* cultures. Alan Hall and Chris van der Merwe for their assistance with the Electron Microscopy of bacteriophages. Renate Filter for automated sequencing. Charlotte Ellis for kindly providing native BoNT and anti-sera and for assistance with Western blot analysis. Pam de Waal and Michelle van Niekerk for advice and help with tissue culture work.

Personnel and students in the Department of Microbiology and Plant Pathology of the University of Pretoria for their interest, support and advice during this study. Especially the students in the Laboratory for Virology and Molecular microbiology for their support, good will, and valuable advice.

Nicci Page, Frederik Hammes, Wilma Havenga, Cornelia Gull, and Elizabeth de Jager for their valued friendship, encouragement and interest in this study.

My family-in-law for their support and interest.

My parents, who encouraged and provided me with the opportunity to commence my tertiary education and for their consistent support throughout my studies.

My husband, André, for his love, patience, and encouragement to reach for my dreams.

SUMMARY

MOLECULAR CLONING AND EXPRESSION OF TYPE C AND D NEUROTOXIN GENES OF *Clostridium botulinum*

by

JENNIFER ROSSOUW

Supervisor: Prof. L. H. Nel
Department of Microbiology and Plant Pathology
University of Pretoria

Co-supervisor: Dr. J. Theron
Department of Microbiology and Plant Pathology
University of Pretoria

for the degree M.Sc

The neuroparalytic syndrome called botulism is caused by the neurotoxins produced by bacteria in the genus *Clostridium*. There are seven toxigenic types of *C. botulinum* (A to G) based on antigenically distinct toxins produced by different strains of the organism. Animal botulism is caused by *C. botulinum* type C and D neurotoxins and has a severe economic impact on cattle farming in South Africa and neighbouring countries. Current treatment regimes include the use of acetylcholine for symptomatic treatment, but this is unfortunately very seldom successful. All indications are that there is no cure for this disease and that effective control can only be achieved through development of efficacious vaccines. The botulinum vaccine currently in use in South Africa contains an adjuvanted toxoid form of the type C and D neurotoxins. However, this bivalent vaccine relies on problematic anaerobic cultivation of the *Clostridium* bacterium followed by isolation, purification and inactivation of the toxin by treatment with formalin. Apart from the fastidious growth requirements of this organism, it has been reported that the production of toxin by these cells declines rapidly and eventually ceases, following laboratory passaging of the bacterial cultures. In addition, improper inactivation of the toxins may also lead to the demise of animals following vaccination. Thus, there exists a great need for a safe, effective and inexpensive vaccine against botulism.

To investigate the potential of types C and D botulinum neurotoxins as efficacious recombinant vaccine candidates against botulism, full-length copies of the genes were obtained by polymerase chain reaction (PCR) amplification from bacteriophage DNA isolated from *Clostridium botulinum* type C (Stockholm) and D (South Africa) cultures. The full-length genes were cloned and subsequently sequenced to verify their integrity. By making use of PCR-based site-directed mutagenesis procedures, three amino acid mutations were introduced in the zinc-binding motif of the respective neurotoxins. Mutation of this domain has previously been reported to successfully detoxify type C neurotoxin. The wild-type and mutant genes were subsequently expressed in insect cells using the BAC-to-BAC™ baculovirus system. Although, unique protein bands corresponding to the size of the neurotoxins could not be seen in Coomassie brilliant blue-stained gels, Western blot analysis indicated immunoreactive material for wild-type and mutant type C corresponding to the size of the type C toxin light chain. However, there was no conclusive evidence to support the successful expression of the full-length wild-type and mutant type D genes.

CONTENTS

PAGE

ACKNOWLEDGEMENTS ... i
SUMMARY ... ii
CONTENTS ... iv
LIST OF ABBREVIATIONS ... viii
LIST OF FIGURES ... xi
LIST OF TABLES ... xiv

CHAPTER 1

LITERATURE REVIEW ... 1
1.1 GENERAL INTRODUCTION ... 1
1.2 HISTORY OF BOTULISM ... 3
1.3 BOTULISM IN LIVESTOCK ... 5
1.3.1 Aetiology ... 5
1.3.2 Epidemiology ... 6
1.3.3 Clinical symptoms ... 7
1.3.4 Diagnosis ... 7
1.4 BOTULINUM NEUROTOXIN ... 8
1.4.1 Location and characterization of neurotoxin genes ... 8
1.4.2 Neurotoxin structural features ... 9
1.4.3 Mechanism of action ... 12
1.4.3.1 Cell binding ... 12
1.4.3.2 Internalization ... 12
1.4.3.3 Membrane translocation ... 12
1.4.3.4 Target modification in cytosol ... 13
1.4.4 Botulinum toxin complex ... 14
1.5 BOTULINUM BACTERIOPHAGES ... 14
1.6 EXPRESSION SYSTEMS ... 16
1.6.1 Prokaryotic expression systems ... 16
1.6.1.1 Expression in *Escherichia coli* ... 16
1.6.1.2 Expression in *Bacillus* ... 17
1.6.2 Eukaryotic expression systems ... 17
1.6.2.1 Baculovirus expression system ... 17
1.6.2.2 Yeast expression systems ... 19

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

1.7 OBJECTIVES OF THE STUDY.. 20

CHAPTER 2

CLONING AND SITE-DIRECTED MUTAGENESIS OF NEUROTOXIN GENES OF *Clostridium botulinum* TYPES C AND D

2.1 INTRODUCTION..21

2.2 MATERIALS AND METHODS..22

2.2.1 Materials..22

2.2.2 Transmission Electron Microscopy (TEM) of bacteriophages.........25

2.2.3 Isolation of bacteriophage DNA..25

2.2.4 Agarose gel electrophoresis..25

2.2.5 Amplification of botulinum neurotoxin genes.............................26

2.2.6 Purification of DNA fragments from agarose gels........................26

2.2.7 Cloning of amplicons into the pGEM®-T Easy vector.................. 26

2.2.7.1 Ligation of amplicons and plasmid vectors.................... 26

2.2.7.2 Preparation of competent cells.......................................27

2.2.7.3 Transformation of competent cells..................................27

2.2.8 Plasmid DNA extraction...28

2.2.9 Restriction enzyme digestion..28

2.2.10 Construction and cloning of mutant botulinum neurotoxin genes....29

2.2.10.1 Inverse PCR...29

2.2.10.2 Construction of clones containing mutated neurotoxin genes.. 29

2.2.11 Automated sequencing of selected clones 30

2.2.11.1 Purification of plasmid DNA.......................................30

2.2.11.2 Automated sequencing...30

2.3 RESULTS.. 31

2.3.1 Construction of clones containing the full-length copies of the type C and D neurotoxin genes..31

2.3.2 Sequencing of full-length botulinum neurotoxin genes 37

2.3.3 Construction of mutant clones... 47

2.4 DISCUSSION.. 51

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

CHAPTER 3

EXPRESSION OF WILD-TYPE AND MUTANT NEUROTOXIN GENES OF *Clostridium botulinum* TYPE C AND D IN THE BAC-TO-BAC™ BACULOVIRUS EXPRESSION SYSTEM

3.1 INTRODUCTION......54
3.2 MATERIALS AND METHODS......57
3.2.1 Materials......57
3.2.2 Construction and characterization of recombinant donor plasmids. 57
3.2.2.1 Restriction enzyme digestions......57
3.2.2.2 Purification of DNA fragments from agarose gels......58
3.2.2.3 Ligation of DNA fragments......58
3.2.2.4 Preparation of competent cells......58
3.2.2.5 Transformation of competent cells......59
3.2.2.6 Screening of recombinant clones......59
3.2.2 Construction of recombinant Bacmid DNA......60
3.2.3.1 Preparation of competent *E. coli* DH10Bac™ cells......60
3.2.3.2 Transformation of competent *E. coli* DH10Bac™ cells......60
3.2.3.3 Isolation of recombinant Bacmid DNA......61
3.2.4 Transfection of *S. frugiperda* cells......61
3.2.5 Titration and purification of baculovirus stocks......62
3.2.6 Analysis of recombinant baculoviruses......63
3.2.6.1 Bacmid DNA preparation......63
3.2.6.2 Selection of primers......63
3.2.6.3 Heminested PCR......64
3.2.7 Fractionation of infected *S. frugiperda* cells......64
3.2.8 SDS-PAGE analysis of proteins synthesized in infected *S. frugiperda* cells......65
3.2.9 Immunological detection of expressed proteins......65
3.2.9.1 Preparation of acetone powders and adsorption of anti-sera......65
3.2.9.2 Western blot analysis......66
3.3 RESULTS......67
3.3.1 Construction of donor plasmids containing the wild-type and mutant type C and D neurotoxin genes......67
3.3.2 Construction of recombinant baculoviruses......72
3.3.3 Analysis of proteins synthesized in infected *S. frugiperda* cells......75

3.4 DISCUSSION 80

CHAPTER 4
CONCLUDING REMARKS 84

REFERENCES 88

LIST OF ABBREVIATIONS

α/β SNAP	soluble NSF accessory protein
°C	degrees Celsius
μg	microgram
μl	microliter
ρmol	picomole
A/a	adenine
aa	amino acid
AcNPV	*Autographa califomica* nuclear polyhedrosis virus
Ala	alanine
Amp^R	ampicillin resistance
Arg	arganine
Asn	asparagine
ATP	adenosine-5'-triphosphate
BoNT	botulinum neurotoxin
BoNT/C	Botulinum neurotoxin type C
BoNT/D	botulinum neurotoxin type D
bp	basepair
BSA	bovine serum albumin
C/c	cytosine
ca.	approximately
cm^2	square centimeters
ddT	dideoxythymidine
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside-5'-triphosphate
ds	double stranded
DTT	dithiothreitol
e.g.	*exempli gratia* (for example)
EDTA	ethylenediaminetetra-acetic acid
et al.	*et alii* (and others)
FCS	foetal calf serum
Fig.	figure
G/g	guanine
Glu	glutamate

Gly	glysine
Gm^R	gentamycin resistance
H	heavy
h	hour
Hc	50 kDa carboxyl-terminal half of the heavy chain of BoNT/TeNT
His	histidine
H_N	50 kDa amino-terminal half of the heavy chain of BoNT/TeNT
Ile	Isoleucine
IPTG	isopropyl β-D-thiogalactosidase
kb	kilobase pairs
kDa	kilodalton
KPTA	potassium phosphate tungstic acid
kV	kilovolts
L	light
LB-medium	Luria-Bertani medium
LD_{50}	lethal dose to kill 50% of experimental organisms
Leu	leucine
M	molar
mA	milliampere
Met	methionine
mg	milligram
min	minute
ml	millilitre
mM	millimolar
mm	millimetre
MOI	multiplicity of infection
mRNA	messenger RNA
MW	molecular weight
ng	nanogram
nm	nanometer
NMJ	neuromuscular junction
NSF	N-ethyl-maleimide-sensitive protein
nt	nucleotide
OD_{550}	optical density at 550 nanometres
OD_{600}	optical density at 600 nanometres
PAGE	polyacrylamide gel electrophoresis
PBS	protein buffered saline

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque forming units
PMSF	phenylmethylsulfonyl fluoride
pPolh	polyhedrin promoter
pro	proline
PSB	protein solvent buffer
RNA	ribonucleic acid
rpm	revolutions per minute
s	second
SDS	sodium dodecyl sulphate
Ser	serome
Sf	*Spidoptera frungiperda*
SFM	serum free medium
SNAP-25	synaptosomal associated protein of 25 kDa
T/t	thymine
TEM	transmission electron microscopy
TEMED	N,N,N^1,N^1-tetramethylethylendiamin
TeNT	Tetanus neurotoxin
Thr	threonine
Tn7L	transposable element 7 left end
Tn7R	transposable element 7 right end
Tris	tris-hydroxymethyl-aminomethane
U	units
UHQ	ultra high quality
UV	ultraviolet
V	volts
v/v	volume per volume
Val	valine
VAMP	vesicle-associated membrane protein
w/v	weight per volume
X-gal	5-bromo-4-chloro-3-indolyl β-D-galactopyranoside

LIST OF FIGURES

PAGE

Fig. 1.1 Mosaic type structure of neurotoxins produced by *C. botulinum* types C and D....................10

Fig. 1.2 Structure-function relationship of botulinum neurotoxins....................11

Fig. 1.3 Target and cleavage sites of botulinum neurotoxins within the neuroexocytosis apparatus....................13

Fig. 2.1 Schematic presentation of specific primer annealing positions and direction of sequencing, in the BoNT/C_{ST} gene and adjacent pGEM®-T Easy vector.................... 24

Fig. 2.2 Schematic presentation of specific primer annealing positions and direction of sequencing, in the BoNT/D_{SA} gene and adjacent pGEM®-T Easy vector.................... 24

Fig. 2.3 (A) Transmission electron micrographs of bacteriophages from *Clostridium botulinum* type C_{ST}(Stockholm) cultures stained with KPTA.................... 32

Fig. 2.3 (B) Transmission electron micrographs of bacteriophages from *Clostridium botulinum* type D_{SA}(South Africa) cultures stained with KPTA.................... 32

Fig. 2.4 Agarose gel analysis of the 3.8 kb fragment obtained by PCR amplification using bacteriophage DNA as template.................... 34

Fig. 2.5 Partial restriction enzyme map of recombinant pGEM®-T Easy vector containing BoNT/C_{ST} or BoNT/D_{SA} gene....................35

PAGE

Fig. 2.6 Agarose gel electrophoretic analysis of the recombinant plasmids, pGEM-C_{ST} and pGEM-D_{SA} constructed by cloning of the respective BoNT amplicons into linearized pGEM®-T Easy vector................ 36

Fig. 2.7 Comparison of complete nucleotide and deduced amino acid sequence of cloned BoNT/C_{ST} gene with the published sequence.. 39 – 42

Fig. 2.8 Comparison of complete nucleotide and deduced amino acid sequence of cloned BoNT/D_{SA} gene with the published sequence.. 43 – 46

Fig. 2.9 Agarose gel electrophoresis indicating amplification of the pGEM®-T Easy vector together with the BoNT/C_{ST} and BoNT/D_{SA} genes........ 48

Fig. 2.10 Partial restriction enzyme map of recombinant pGEM®-T Easy vector containing the mutated BoNT/C_{ST} or BoNT/D_{SA} gene.................... 49

Fig. 2.11 Agarose gel electrophoretic analysis of the recombinant pGEM®-T Easy vector containing either the C_{ST} or D_{SA} genes, obtained by selfligation of the mutagenized PCR amplicons prior to transformation of competent cells... 50

Fig. 3.1 Partial restriction enzyme map of recombinant pFASTBAC™DUAL vector containing BoNT/C_{ST} or BoNT/MutC_{ST} gene..................... 68

Fig. 3.2 Agarose gel electrophoretic analysis of the recombinant donor plasmids, pFASTBACDUAL-C_{ST} and pFASTBACDUAL-MutC_{ST}, constructed by cloning of the respective BoNT DNA fragments into linearized pFASTBAC™DUAL vector... 69

Fig. 3.3 Partial restriction enzyme map of recombinant pFASTBAC™DUAL vector containing BoNT/D_{SA} or BoNT/MutD_{SA} gene..................... 70

PAGE

Fig. 3.4 Agarose gel electrophoretic analysis of the recombinant donor plasmids, pFASTBACDUAL-D$_{SA}$, and pFASTBACDUAL-MutD$_{SA}$ constructed by cloning of the respective BoNT DNA fragments into linearized pFASTBAC™DUAL vector.. 71

Fig. 3.5 Agarose gel electrophoresis of amplicons obtained by semi-nested PCR for verification of the presence of BoNT/C and BoNT/MC genes in the recombinant Bacmid DNA.. 73

Fig. 3.6 Agarose gel electrophoresis of amplicons obtained by semi-nested PCR for verification of the presence of BoNT/D and BoNT/MD genes in the recombinant Bacmid DNA.. 74

Fig. 3.7 SDS-PAGE gel electrophoresis of recombinant baculovirus-infected *S. frugiperda* cells under reducing conditions in the presence of 2-mercaptoethanol.. 77

Fig. 3.8 Western blot analysis of Baculo-BoNT/C and Baculo-BoNT/MC-infected *S. frugiperda* cell extracts subjected to SDS-PAGE electrophoresis under reducing conditions in the presence of 2-mercaptoethanol.. 78

Fig. 3.9 Western blot analysis of Baculo-BoNT/D and Baculo-BoNT/MD-infected *S. frugiperda* cell extracts subjected to SDS-PAGE electrophoresis under reducing conditions in the presence of 2-mercaptoethanol.. 79

LIST OF TABLES

PAGE

Table 1.1 Botulinum neurotoxin producing *Clostridium* species.......... 6

Table 2.1 Sequences of PCR primers used for amplification of botulinum neurotoxin genes.......... 23

Table 2.2 Sequences of primers used for sequencing of full-length botulinum neurotoxin genes.......... 23

Table 2.3 Amino acid differences between cloned BoNT/C_{ST} and published sequence.......... 38

Table 2.4 Amino acid differences between cloned BoNT/D_{SA} and published sequence.......... 38

Table 2.5 Nucleotide and derived amino acid sequences of zinc-binding motif of wild-type and mutant type C_{ST} and D_{SA} botulinum neurotoxin genes.......... 51

CHAPTER 1

LITERATURE REVIEW

1.1 GENERAL INTRODUCTION

"Disease can never be conquered, can never be quelled by emotion's wailful screaming or faith's cymballic prayer. It can only be conquered by the energy of humanity and the cunning in the mind of man. In the patience of a Curie, in the enlightenment of a Faraday, a Rutherford, a Pasteur, a Nightingale, and all other apostles of light and cleanliness, rather than of a woebegone godliness, we shall find final deliverance from plague, pestilence, and famine." (Sean O'Casey (1884-1964), Isnishfallen, Fare Thee Well, vol.1, 1949).

Can disease really be conquered by the energy of humanity and the cunning in the mind of man or will microbes win the battle? In the microbial world, warfare is a constant and the survival of most organisms necessitates the demise of others. Yet there are times of extraordinary collectivity in the microbial world, when the rivalry is abandoned and forces joined to combat a shared enemy, humankind. Microbes are evolving far more rapidly than humans or animals, adapting to changes in their environments by mutating, undergoing high-speed natural selection, or drawing plasmids and transposons from the vast genetic pool in their environments. The most sophisticated of their species have the ability to outwit or manipulate the only microbial sensing system humans and animals possess, their immune system.

Two centuries have past since Edward Jenner first realized that the host immune system could be used to outsmart disease-causing organisms. The cells of the immune system are taught about the possible attack of microbes or their toxins. Immunization is currently one of the most effective and versatile means to combat infectious diseases. The first vaccines were the result of keen observation and experimentation and they consisted of live attenuated, killed inactivated microbes or purified microbial components.

The need for immunization agents against botulism was evident as soon as the economic importance of this fatal intoxication was realized. Botulism is caused by seven serologically distinct botulinum neurotoxins (A to G) produced by several *Clostridium* species. Adsorption of the ingested botulinum neurotoxins from the intestinal tract and transport to central nervous system gives rise to a group of symptoms, which involves digestive disturbances, nausea, vomiting, and partial or complete paralysis of the muscles.

Botulism in livestock, commonly associated with types C and D neurotoxins, has a severe economic impact on cattle farming in southern Africa, Australia and South America. Current treatment regimes include the use of acetylcholine for symptomatic treatment, but this is unfortunately very seldom successful. All indications are that there is no cure for this disease and that effective control can only be achieved through development of efficacious vaccines. The first effective animal toxoid was prepared through the detoxification of type C cultures grown in cooked meat medium by 0.8% formalin and precipitated with 0.6% potassium aluminium sulfate (Smith and Sugiyama, 1988). In southern Africa, botulism in livestock is controlled using a commercially available vaccine produced by Onderstepoort Biological Institute. Despite the efficacy of these vaccines, they are very costly to produce. Botulinum neurotoxins are not produced in high levels in culture and obtaining sufficient amounts of toxin for purification and toxoiding is laborious (Clayton *et al.,* 1995). Another problem commonly experienced is that production of toxin by cell cultures declines rapidly and eventually ceases, following passaging of the bacterial culture.

Recombinant biotechnology may provide solutions to problems regarding botulinum vaccine production. A shift from classic vaccines to recombinant and sub-unit vaccines, with the aim of developing a cheaper and more efficacious botulinum vaccine, is evident. This technology allows for genes to be transferred and expressed as recombinant products in easily grown organisms such as *Escherichia coli*, *Bacillus subtilis*, yeasts, and insect cells.

This chapter reviews several aspects regarding botulism, as well as some of the prokaryotic and eukaryotic expression systems commonly used in biotechnology.

1.2 HISTORY OF BOTULISM

The story of botulism starts as early as 1735 in Germany where the symptoms of botulism were first described (Viljoen, 1965). During the remainder of the eighteenth century cases of botulism were reported sporadically from various parts of Europe, Russia, England and America (Viljoen, 1965 and references therein). It was the well-recorded outbreak of 1793 in Wildbad, Wurttemburg, which really prompted serious study of botulism (Smith and Sugiyama, 1988). At the time two opinions were advanced as to the cause of the illness of thirteen people, and the subsequent death of six. According to the chief medical officer of Wurttemburg, the outbreak was due to belladonna poisoning, but a more widely accepted view was that the outbreak was caused by the consumption of *Bluzen* or *Schweinsmagen*, a type of blood sausage (Smith and Sugiyama, 1988).

During the same period Le Vaillant also recorded cases of a similar disease in cattle, referred to as *Lamsiekte*. He considered *Lamsiekte* to be one of the four most destructive diseases of cattle in the Cape region at the time (Henning, 1949). Commissioner de Mist (1805) and Lichtenstein (1803-6) also recorded the occurrence of *Lamsiekte* in several parts of the Cape (as referenced in Henning, 1949). During the nineteenth century, *Lamsiekte* prevailed and caused great losses of cattle in Southern Africa.

The increase in the number of reported cases of sausage poisoning during the next decades drew the attention of the district medical officer, Justinus Kerner (1786-1862). Kerner made an extensive study of the disease and described botulism as "a special kind of poisoning, quite unlike any other" (Smith and Sugiyama, 1988). However, it was not until 1896 that Emile P.M. van Ermengem isolated the causative organism from the remains of raw, salted ham which had caused neuromuscular paralysis among a group of musicians in the small Belgian village of Ellezelles (Viljoen, 1965; Smith and Sugiyama, 1988). Van Ermengem considered that the Ellezelles patients had the same disease as sausage poisoning and suggested that this spore-forming obligatory anaerobic bacterium be called *Bacillus botulinus* (Smith and Sugiyama, 1988).

Early in the twentieth century an interesting discovery was made from a can of white beans. In 1904 Landmann isolated a second type of organism producing a toxin resembling that of the organism isolated by Van Ermengem (Smith and Sugiyama, 1988). In 1917, the Society of American Bacteriologists' Committee on Classification of Bacterial types recommended the generic term Clostridium as suggested by Zopf for this group of organisms (Viljoen, 1965). Burke established that more than one type of organism is responsible for the disease and that these organisms each produce a different type of toxin. This prompted the designation of the organisms isolated by Van Ermengem and Landmann as type B and A, respectively (Burke, 1919).

Three years later type C was isolated, almost simultaneously in Australia and in the United States. In the US, Ida Bengston isolated strains of clostridia, from green fly larvae (*Lucilia caesar*), which is involved in *Limbemeck* among fowls. She found that the toxin they produced was serologically unrelated to that of types A and B (Bengston, 1922). Seddon isolated a related organism from decomposed bone found in a part of Australia where "Tasmanian Midland Disease" occurred (Seddon, 1922). He called the causative organism *Bacillus parabotulinus*. Gunnison and Meyer (1929) later designated the two type C strains as C-alpha (Bengston strain) and C-beta (Seddon strain) (Kriek and Odendaal, 1994).

In 1919, Theiler and his collaborators proved that *Lamsiekte* in cattle was caused by ingestion of carcass material picked from the veld (Kriek and Odendaal, 1994). It was established that the toxin is elaborated by a specific anaerobic bacterium that grows luxuriantly in carcass debris. *C. botulinum* type D was finally isolated in South Africa by Du Toit and Robinson of the Veterinary Research Institute at Onderstepoort, from the carcass of a cow that had died of *Lamsiekte*. Robinson later stated that, although type D toxin is mainly responsible for *Lamsiekte*, type C also exists on farms and should not be excluded as a cause of this hazard in South Africa (Kriek and Odendaal, 1994).

Type E was encountered simultaneously during the mid 1930's in Russia and the United States. Subsequently, Bier isolated type E from spoiled fish that was a source of intoxication (Smith and Sugiyama, 1988). In 1986 another species of *Clostridium* was implicated in causing botulism when McCroskey and co-workers isolated a strain of *C. butyricum* from an infant with type E botulism (McCroskey *et al.*, 1986).

Moller and Scheibel (1960) isolated type F from home-made liver paste that was implicated in an outbreak of botulism on the Danish Island of Langeland. In 1985, Hall and co-workers implicated a strain of *C. barati*, which produce type F botulinum neurotoxin, as the cause of botulism in infants (Hall *et al.*, 1985).

Giménez and Ciccarelli (1970) isolated type G from soil of a cornfield in Mendoza, Argentina. Type G has been associated with sudden unexpected deaths in Argentina (Sonnabend *et al.*, 1981), but has not been clearly implicated as the cause of paralytic illness or death (Suen *et al.*, 1988). Suen and co-workers (1988) proposed that type G be renamed *Clostridium argentinense* on the grounds of its metabolic and cultural differences.

1.3 BOTULISM IN LIVESTOCK

1.3.1 Aetiology

Botulinum neurotoxins of *C. botulinum* types C and D appear to be the exclusive cause of botulism of livestock in countries in the southern hemisphere, such as South Africa, Australia, and certain South American countries. However, neurotoxins from type B are more often the cause of botulism in livestock in the USA and Europe.

The botulinum neurotoxin-producing clostridium is a diverse group of organisms originally grouped together on the basis of the toxin that they produce. However, it has been shown that botulinum toxin production is a poor parameter on which to base species classification and identification (Suen *et al.*, 1988). Currently, these clostridia can be divided into six groups. These groups are very different in their phenotype, but members of two or more groups may share the same toxin. Group one contains proteolytic strains of *C. botulinum* which produce toxin type A, B, or F. Strains from this group is the most able to colonize the digestive tract. Group two contains non-proteolytic strains of *C. botulinum* which produce toxin types B, E or F. The third group is weakly proteolytic or non-proteolytic *C. botulinum* strains that produce toxin types C1, C2, or D. The fourth group contains the proteolytic but non-saccharolytic *C. argentinense* strains that produce type G toxin (Suen *et al.*, 1988). Group five consists

of proteolytic *C. butyricum* that produces type E toxin, while group six contains non-proteolytic *C. barati* that produces type F toxin.

Table 1.1 Botulinum neurotoxin-producing *Clostridium* species

Group	Strain	Toxin type	Cultural characteristic
Group I	*C. botulinum*	A, B, F	Proteolytic
Group II	*C. botulinum*	B, E, F	Non-proteolytic
Group III	*C. botulinum*	C, D	Weak or Non-proteolytic
Group IV	*C. argentinense*	G	Proteolytic, but Non-saccharolytic
Group V	*C. butyricum*	E	Proteolytic
Group VI	*C. barati*	F	Non-proteolytic

The botulinum neurotoxin-producing clostridia are anaerobic, gram-positive, straight to slightly curved rods (2 – 15 x 0.5 – 2 μm). They loose their staining ability and old cells tend to stain gram-negative. All the strains produce oval and subterminal spores, except *C. argentinense* for which spores can either be present or absent.

1.3.2 Epidemiology

Botulism has been described in cattle, horses and foals, mules and donkeys, sheep and goat, pigs, water birds, turkeys, ostriches and ducks, and dogs (Kriek and Odendaal, 1994 and references therein). Outbreaks in livestock can generally be subdivided into two main categories, namely those associated with phosphorus deficiency and osteophagia (bone craving), and secondly the form that is associated with the ingestion of toxic feed or water.

Botulism associated with phosphorus deficiency and osteophagia occurs sporadically over a relatively long period of time in mainly cattle, but also to a lesser extent in sheep and goats. Cattle suffering from osteophagia contract the disease by ingesting toxic material lying in pastures. They most commonly develop the subacute to chronic form of the disease that manifest between 7 and 10 days after toxin ingestion. A few may even be so mildly affected that they escape notice.

Botulism in livestock which is caused by the ingestion of toxic feed or water may result in the deaths of large numbers of animals over a short period of time (ca. 24 h) and new clinical cases are detected over a period of days. At the beginning of the outbreak, the cattle usually die acutely and during the outbreak the duration of the clinical symptoms increase.

Toxico-infectious botulism is a rare form of the disease in humans and animals. In this instance, the toxin is produced *in vivo* by *C. botulinum* growing in the body, rather than ingestion of the toxin (Smith and Sugiyama, 1988). It may either follow wound infection or occur after the ingestion of spores and the subsequent germination thereof in the gastrointestinal tract.

1.3.3 Clinical symptoms

The clinical symptoms in cattle vary considerably and depend on the quantity of toxin that has been ingested. The disease can be divided into peracute, acute, subacute and chronic forms, based upon the onset, the sequence of development of paralysis and severity of the clinical signs (Henning, 1949). Affected animals are afebrile and typically manifest partial or complete flaccid paralysis of the muscles of locomotion, mastication and deglutition. Paralysis of muscles usually commences in the hindquarters and then spreads progressively forwards, affecting in turn the forelimbs, the neck and head.

1.3.4 Diagnosis

Diagnosis of botulism is often difficult and a presumptive diagnosis is usually made on the basis of history, clinical signs and negative post-mortem examination. The systems for the detection of botulinum neurotoxin in specimens taken from diseased or dead

animals lack sensitivity and it is often difficult to confirm the diagnosis of botulism. Methods that are used include mouse bioassays, immunodiffusion, radio-immuno-assays, various immunochemical assays, and temperature-induced microcomplement fixation tests (Kriek and Odendaal, 1994 and references therein).

The possibility of other diseases causing similar clinical symptoms should be excluded. In South Africa the two diseases that are most commonly confused with botulism are diplodiosis (mycotoxicosis caused by *Diplodia maydis* poisoning) in cattle and krimpsiekte (plant poisoning caused by members of the Crassulaceae containing bufadienolides) in sheep and goats. In cattle other diseases that mimic some of the clinical signs of botulism include: ingestion of the toxin of the fungus *Aspergillus clavatus*, organophosphate and lead poisoning, and infectious diseases such as rabies, listeriosis, bovine ephemeral fever, sporadic bovine encephalomyeltis, thrombotic meningoencephalitis, cerebral babesiosis and cerebral theileiriosis, as well as non-specific bacterial meningitis and ecephalitides. In sheep the following should also be considered, tick paralysis caused either by *Oxodes rubicundus* or *Rhipicephalus evertsi*, focal symmetrical encephalomalacia, twin lamb disease (domsiekte) and *Coenurus cerebralis* infestation.

1.4 BOTULINUM NEUROTOXIN

Botulinum neurotoxins are a family of seven structurally similar, but antigenically different proteins produced by different strains of *Clostridium botulinum, Clostridium barati, Clostridium butyricum* and *Clostridium argentinense* (Hall *et al.*, 1985; McCroskey *et al.*, 1986; Suen *et al.*, 1988). These proteins are the most toxic substance known to man (LD_{50} of 0.1 – 1.0 ng/kg) and specifically function on the neuro-muscular junctions to block neurotransmitter release (Montecucco and Schiavo, 1994).

1.4.1 Location and characterization of neurotoxin genes

The location of the respective botulinum neurotoxin genes is not yet clear. Several locations have been investigated, but only the type C and D genes has been experimentally proven to be of an extrachromosomal bacteriophage origin. Other

options investigated is plasmid- (type G) and chromosomally-encoded genes (A, B, E, and F) (Minton, 1995).

The phage origin of the neurotoxin genes was demonstrated when both the BoNT/C1 and BoNT/D genes were cloned from their respective bacteriophages (Binz *et al.*, 1990; Hauser *et al.*, 1990; Kimura *et al.*, 1990). The obtained nucleotide sequences of the genes encoding botulinum neurotoxin type C_1 from *C. botulinum* type C strains 468 and Stockholm show great sequence similarity (99.9% homology). The gene encoding botulinum neurotoxin type D isolated from type D strains 1873 and CB-16 are also similar in sequence (99.9% homology) (Moriishi *et al.*, 1996(b)). However, it was found that the neurotoxin produced by some type C and D strains are not completely consistent with BoNT/C1 and BoNT/D on the basis of antigenicity (Moriishi *et al.*, 1989; Ochanda *et al.*, 1984). The type C neurotoxin isolated from strain 6813 has antigenic structures from the BoNT/C1 light chain and the BoNT/D heavy chain. The amino-terminal region of C6813 (Met-1 to Leu-526) display a 96% sequence homology with type C1 (Met-1 to Leu-526), while the carboxy-terminal region of C6813 (Asn-902 to Glu-1280) display a 96% homology with that of type D (Asn-898 to Glu-1276) (Moriishi *et al.*, 1996(b)). On the other hand, the type D neurotoxin isolated from the South African strain has antigenic structures from the BoNT/D light chain and the BoNT/C1 heavy chain. The amino-terminal region of D_{SA} (Met-1 to Val-522) displays a 96% homology with type D (Met-1 to Asn-521), while the carboxy-terminal region of D_{SA} (Trp-945 to Glu-1285) displays a 74% homology with that of type C1 (Trp-949 to Glu-1291) (Moriishi *et al.*, 1996(a)). This indicates that the type C and D neurotoxins consist of a mosaic-type structure and that at least 4 isoforms can be distinguished (Fig. 1.1).

1.4.2 Neurotoxin structural features

The botulinum neurotoxins are synthesized intracellularly as 150 kDa single-chain polypeptides that are released into the culture supernatant through autolysis of the cells after completion of growth cycles. Endogenous or exogenous proteases nick the chain within an exposed loop to generate the active di-chain neurotoxins composed of a heavy chain (H, 100 kDa) and a light chain (L, 50 kDa) bridged by a disulphide bond.

The botulinum neurotoxins are folded into three functionally distinct domains (L, H_N, and H_C) (Fig. 1.2). The L domain is responsible for zinc-dependent proteolysis of protein components of the neuroexocytosis apparatus. The botulinum neurotoxins bind zinc via two histidines and a glutamic acid located in a zinc-binding motif (His-Glu-aa-aa-His) in the middle of the L domain. The H chain can be divided into two domains (H_C and H_N) corresponding to the carboxy- and amino-terminal halves of the H chain, respectively. The H_C domain seems to be mainly responsible for neurospecific binding to receptors on the surface of the presynaptic membrane, while the H_N domain appears to govern internalization and membrane translocation.

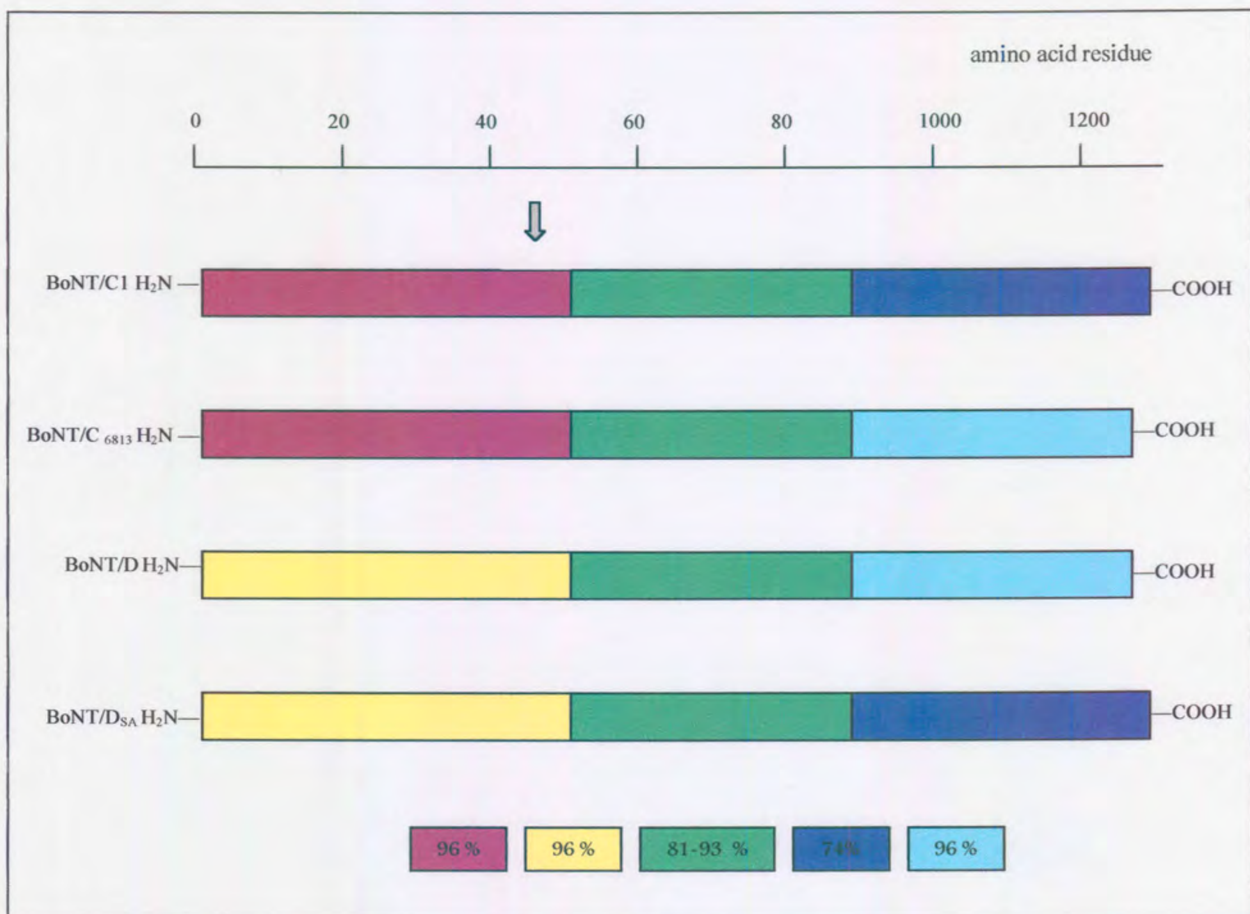

Fig. 1.1 Mosaic-type structure of neurotoxins produced by *C. botulinum* types C and D (Moriishi *et al.*, 1996a,b). Nicking sites for disulfide bond formation are located by the arrow. Highly homologous regions are illustrated in similar colours.

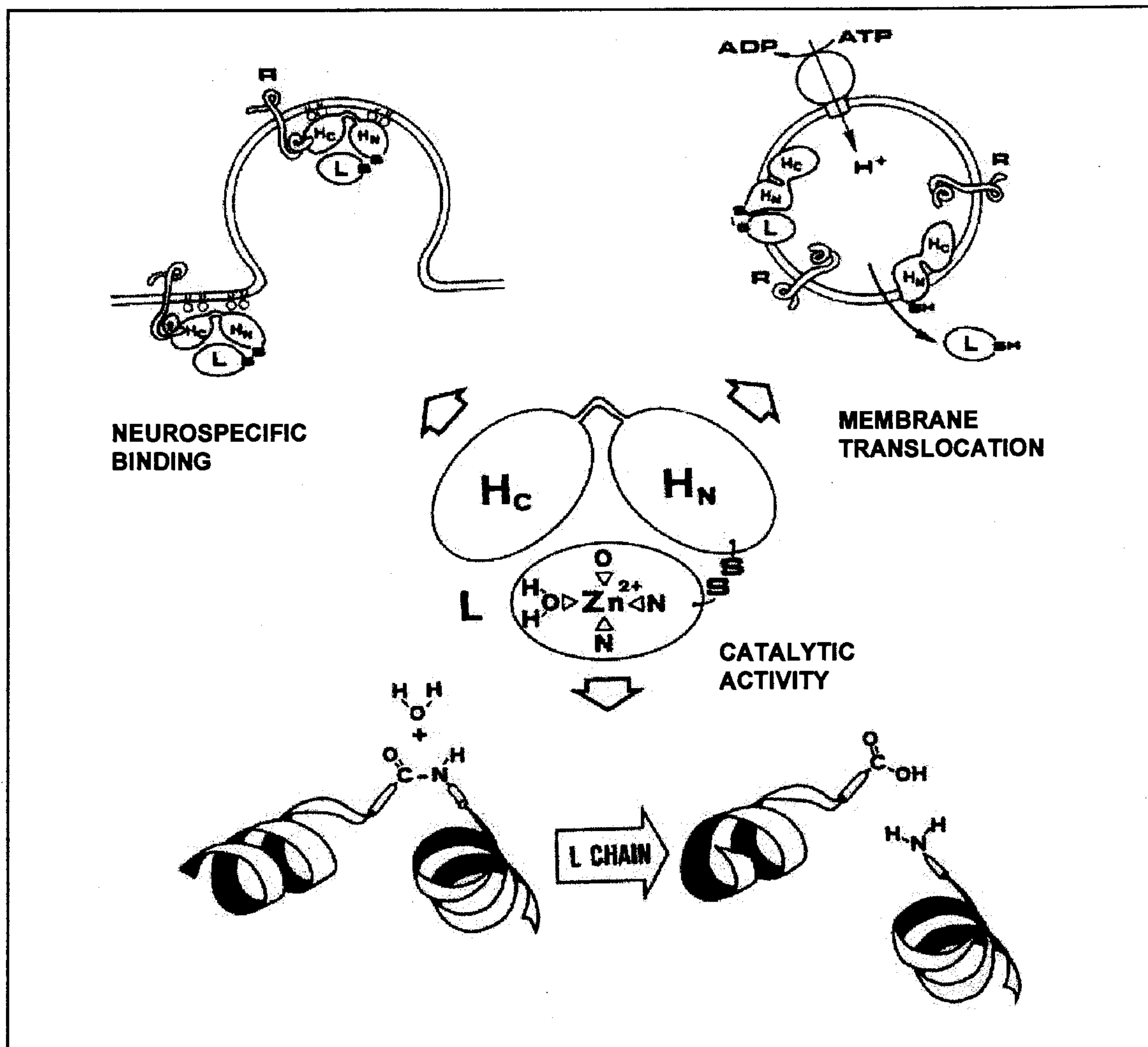

Fig. 1.2 Structure-function relationship of botulinum neurotoxins (from Montecucco and Schiavo, 1994). The 50 kDa carboxy-terminal domain of the H chain (H_C) of botulinum neurotoxins is mainly responsible for neurospecific binding to receptors on the surface of the presynaptic membrane (top-left panel). Binding is followed by the internalization of the toxin-receptor complex inside vesicles, which remain within the neuromuscular junction. The amino-terminal domain (H_N) is suggested to be involved in membrane translocation of the L chain into the nerve-cell cytoplasm (top-right panel). The L chain is released into the cytosol by reduction of the interchain disulphide bond, where it displays zinc-endopeptidase activity by the selective cleavage of proteins of the neuroexocytosis machinery (lower panel).

1.4.3 Mechanism of action

Botulism is essentially a flaccid paralysis that results from the toxin acting on the peripheral instead of the central nervous system. The method whereby cell intoxication occurs, consists of four steps: (i) cell binding, (ii) internalization, (iii) membrane translocation, and (iv) target modification in cytosol (Montecucco and Schiavo, 1994).

1.4.3.1 Cell binding

Botulinum neurotoxins bind specifically to receptors on the surface of nerve cells. The receptors to which the neurotoxins bind is unknown at this stage and may not be the same for all the toxin types. However, it is clear that the receptors are located on the motor neuron plasmalemma at the neuromuscular junction. These receptors should display high affinity to bind the minute concentrations of neurotoxins sufficient to cause death. Various reports indicate that clostridial neurotoxin binding to nerve cells are facilitated by a double-receptor model that generates very high affinity binding (Montecucco and Schiavo, 1994). The neurotoxins are supposed to bind first to the large negatively-charged surface of the presynaptic membrane, which consists mainly of polysialogangliosides and other acidic lipids.

1.4.3.2 Internalization

Although binding is rapid, toxins do not immediately cause observable paralysis of the neuromuscular junction (NMJ). During this period, the neurotoxins are internalized inside vesicles of unknown nature in a temperature- and energy-dependent process (Montecucco and Schiavo, 1994). At this stage, specific anti-sera cannot neutralize the toxin any longer.

1.4.3.3 Membrane translocation

Membrane translocation follows internalization, but there appears to be no coupling between the two steps. The target of the neurotoxins is cytosolic and it is thus important to translocate at least the L catalytic domain across the vesicle membrane. It has been proposed that translocation is triggered by the acidification of the vesicle lumen by a proton-pumping ATPase which leads to conformational change of the toxin. In its acid conformation, the toxin inserts into the lipid bylayer and H_N promotes the translocation of the L chain across the vesicular membrane (Montecucco and Schiavo, 1994).

1.4.3.4 Target modification in cytosol

The intracellular activity of clostridial neurotoxins is a zinc-dependent proteolysis of specific components of the neuroexocytosis apparatus (Fig. 1.3) (Montecucco and Schiavo, 1994). Botulinum neurotoxins of types B, D, F and G recognize and cleave a common intraneural target within the neuroexocytosis apparatus known as VAMP/synaptobrevin (VAMP, vesicle-associated membrane protein). BoNT/A and BoNT/E cause the specific hydrolysis of SNAP-25 (synaptosomal associated protein of 25 kDa), while serotype C cleaves syntaxin.

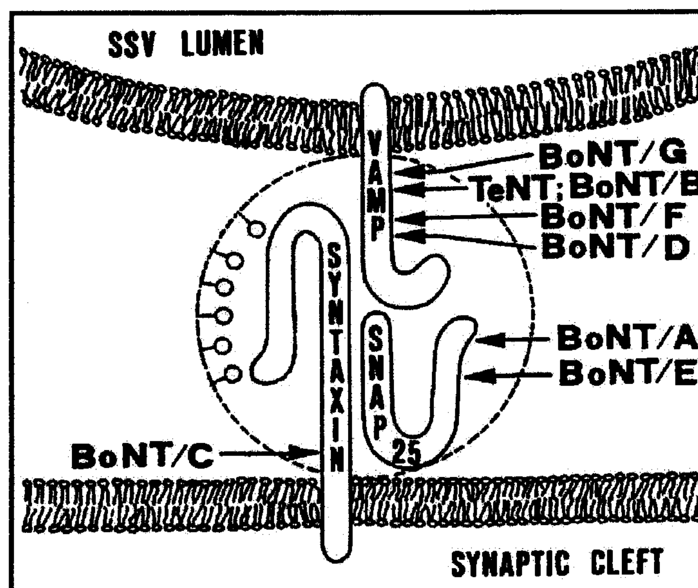

Fig. 1.3 Target and cleavage sites of botulinum neurotoxins within the neuroexocytosis apparatus (from Montecucco and Schiavo, 1994). The L chain displays zinc-endopeptidase activity specific to components of the 20S multi-subunit vesicle docking/fusion complex. Botulinum B, D, F, and G neurotoxins specifically cleave VAMP/synaptobrevin, a protein of the neurotransmitter-containing vesicles. Botulinum neurotoxins types A and E recognize and cleave SNAP-25, a presynaptic membrane component. The type C neurotoxin specifically proteolyses syntaxin which is also localized in the plasma membrane. Arrows indicate the approximate positions of cleavage sites. SSV indicates neurotransmitter containing small synaptic vesicles.

It has been proposed that every cellular event of vesicle docking and fusion with a target membrane is mediated by a complex composed of VAMP, SNAP-25, syntaxin, NSF (N-ethyl-maleimide-sensitive protein), α/β SNAP (soluble NSF accessory protein), γ-SNAP, MUNC-18 and a cysteine-string protein (Montecucco and Schiavo, 1994). Cleavage of components within this complex leads to persistent and sustained blocking of neurotransmitter release.

1.4.4 Botulinum toxin complex

Botulinum neurotoxins are commonly found to be complexed with other, non-toxic proteins and are known as progenitor toxins. Three forms are recognised: M toxin (300 kDa), L toxin (500 kDa) and LL toxin (900 kDa). The M toxin is composed of a BoNT (150 kDa) in association with a similarly sized non-toxic protein (NTNH, 150 kDa). The L and LL progenitor toxins contain an additional undefined number of proteins with haemagglutinin (HA) activity. The nature of progenitor toxin varies between the different toxinogenic types, and more than one form may be produced by a single strain.

Progenitor toxins exhibit greater stability to temperature and pH extremes than do purified botulinum neurotoxins. The non-toxic proteins of the botulinum complex aid the neurotoxin in withstanding the acidic and protease-rich environment of the stomach (Minton, 1995).

1.5 BOTULINUM BACTERIOPHAGES

Bacteriophages in *Clostridium botulinum* were first reported by Vinet *et al.* (1968). Large phages with octahedral heads were observed in electron micrographs taken of cells from type C cultures.

Dolman and Chang (1972) divided the botulinum bacteriophages into four groups. These groups are based on the cultural characteristics of the *Clostridium* host and phage morphology. Group one contains the bacteriophages associated with the proteolytic strains from *C. botulinum* types A, B, and F. This group is characterized by the presence of either or both kinds of phages. The first kind of phage has icosahedral

heads (50 – 75 nm) with contractile tails (85 – 165 nm), while the second phage has bullrushy heads (75 – 85 nm by 35 – 45 nm) with flexible tails (85 – 165 nm).

The second group contains the bacteriophages associated with the non-proteolytic strains of type B and F. Both the type B and F strains contain phages with icosahedral heads (50 – 65 nm) and contractile tails (90 – 110 nm). A different phage associated with type F has octahedral heads (75 – 80 nm) with flexible tails (over 200 nm).

The third group contains the bacteriophages from type C and D that is associated with strikingly large octahedral-headed phages (80 – 100 nm), with long contractile tails (275 – 420 nm) that is usually sheathed. Even though many different strains of *C. botulinum* have been shown to carry bacteriophages, only in the case of BoNT/C and BoNT/D has a direct relationship between bacteriophage and toxicogenicity been established. This relationship was first suggested by Inoue and Iida (1970) who treated non-toxigenic type C mutants with a cell-free lysate of a toxigenic strain and restored toxin production to the mutant. Eklund and Poysky (1974) also demonstrated the interconversion of type C and D strains by specific bacteriophages. They indicated that when type C and D cultures are cured from their phages, they cease to produce their dominant toxin. Furthermore, these cured non-toxic strains could be converted to either toxicogenic type depending on the specific phage used for infection. The ease with which the type C and D clostridial cells loose their phages suggests that the prophages are not stably integrated into the host chromosome. The relationship between the phage and the host organism is one of pseudolysogeny, where the phage exists in an extrachromosomal state, somewhere between virulence and lysogeny. Fujii and co-workers (1988) reported that the type C and D bacteriophages contain double stranded DNA with a molecular weight of 70×10^6 to 100×10^6 (110 - 150 kb).

The fourth group of botulinum bacteriophages from type E has icosahedral heads (50 – 65 nm) with contractile tails (85 – 135 nm) that is occasionally sheathless and often accompanied by surplus tail-like structures.

1.6 EXPRESSION SYSTEMS

The major expression systems that are commercially available to most investigators include *Escherichia coli, Bacillus*, yeasts, mammalian cells and the baculovirus system for expression in insect cells. Each of these systems has its advantages and disadvantages making them suitable for different applications. Recent years have seen an increase in the development of gene expression systems in a variety of other organisms including streptomyces, fungi, pseudomonads, and the industrial gram positive bacteria with low guanine and cytosine content (*Lactobacillus* spp., *Lactococcus* spp., *Stahylococcus* spp., and *Streptococcus* spp.). These expression systems allow for present and future exploitation of these organisms as cell factories for the synthesis of products of medical, agricultural, and food biotechnological importance.

1.6.1 Prokaryotic expression systems

1.6.1.1 Expression in *Escherichia coli*

Escherichia coli has long been the primary prokaryotic host for the synthesis of heterologous proteins. Advanced knowledge of *E. coli* genetics and physiology has accounted for the preferential use of this organism as a host for gene expression. An important advantage of this system is the rapid generation of biomass due to high rates of cell growth and availability of low-cost culture conditions (Balbas and Bolivar, 1990). An important disadvantage of *E. coli* expression is its limited capacity to secrete proteins and its inability to exert certain posttranslational modifications of proteins (e.g. disulfide bond formation, glycosylation, and acetylation), which usually leads to incorrect protein folding and subsequently difficult purification of proteins.

Transformation of *E. coli* is also easy and a wide range of easy-to-use vectors is available. Vectors used for the expression of heterologous proteins should contain several elements to facilitate the expression process. These vectors should have a well-characterized origin of replication, a selection marker for plasmid maintenance, a strong promoter, a cloning region downstream from the promoter in which to insert foreign genes, a ribosome binding site (Shine-Delgarno), and in some cases a translation initiation ATG codon, to ensure transcription initiation and efficient translation (Old and Primrose, 1994). *E. coli* expression vectors often have multiple cloning sites adjacent to

a fusion protein in order to express heterologous proteins as N-terminal or C-terminal fusion proteins. The advantage of fusions is that it stabilizes the proteins from proteolytic degradation, facilitate purification or can be used as a reporter gene to monitor expression. Native protein can be obtained by *in vitro* cleavage of these fusions.

1.6.1.2 Expression in *Bacillus*

The gram-positive organism, *Bacillus,* is well known for it's important contributions to medical, agricultural and food biotechnology (e.g. industrial enzymes, antibiotics, and insecticides). *Bacillus* species, which lack an outer membrane offer a simpler path for protein secretion than *E. coli* and could be considered as alternative hosts when secretion is desired. The proteins need only to cross the cytoplasmic membrane to find themselves in the extracellular medium. However, *B. subtilis* produces a range of endogenous proteases which may probe to be problematic in using this host for foreign protein secretion. Protease-deficient strains have subsequently been generated by mutagenesis, but these proved to be autolytic (Simonen and Palva, 1993; Billman-Jacobe, 1996).

Significant progress has been made in the development of *Bacillus brevis* as an alternative *Bacillus* expression system. *Bacillus brevis* offers high levels of efficient expression and secretion of proteins directly into the culture medium in a soluble and biologically active form, without proteolytic degradation of the secreted proteins in the extracellular milieu (Udaka and Yamagata, 1993).

1.6.2 Eukaryotic expression systems

1.6.2.1 Baculovirus expression system

In recent years the baculovirus expression system has gained widespread use for the production of a wide variety of heterologous proteins (O'Reilly *et al.,* 1992 and references therein). This system provide an eukaryotic environment that is generally conducive to the proper folding, disulfide bond formation, oligomerization, and/or post-translational modification (e.g. glycosylation) required for biological activity of some proteins (O'Reilly *et al.*, 1992). In addition, biologically active proteins are synthesized more reliably than in bacterial or yeast expression systems, and the yields generally exceed those from mammalian expression systems (Liebman *et al.*, 1999).

Baculoviruses are classified in the family *Baculoviridae* and belong to a diverse group of large double-stranded DNA viruses that infect many different species of insects as their natural hosts. The baculovirus genome is replicated and transcribed in the nuclei of infected host cells where the large Baculovirus DNA (between 80 and 200 kb) is packaged into rod-shaped nucleocapsids. Recombinant baculovirus particles can accommodate large amounts of foreign DNA, since the size of the nucleocapsids is flexible. *Autographa californica* nuclear polyhedrosis virus (AcNPV) is the most extensively studied baculovirus. Infectious AcNPV particles enter susceptible insect cells by facilitated endocytosis, and viral DNA is uncoated in the nucleus.

The basic principle of this expression system usually involves foreign genes that are placed under the transcriptional control of the promoters of two dispensable viral gene products, p10 and polyhedrin. These proteins are expressed to very high levels in cells infected with wild-type baculovirus, and their replacement by almost any foreign gene of interest results in heterologous expression to similarly high levels (Davies, 1995). Since the baculovirus genome is generally too large to easily insert foreign genes, heterologous genes are cloned into transfer vectors. The genes are cloned downstream of the required viral promoter and is flanked both sides by viral sequences. The gene and promoter is targeted to a particular region in the viral genome and inserted through *in vivo* homologous recombination.

Recombinants can be produced by making use of two methods that differ only in the location at which the gene of interest is incorporated into the baculovirus genome. In the most common method of recombination, a transfer vector containing the cloned gene is co-transfected with viral DNA into insect cells where the recombination event takes place. Linear forms of the viral genome, which is unable to initiate an infection unless rescued to the circular replication-competent form by recombination, are used (Kits and Possee, 1993). The efficiency of this process relies on the construction of a transfer vector in which the gene of interest is positioned under the control of a strong baculoviral promoter and flanked by baculovirus DNA for recombination. The use of high-quality linear viral DNA with the minimum possibility of recircularization via non-desired recombination events is also important.

The second method of recombinant baculovirus formation involves the introduction of genetic material into the viral genome outside of the insect cells. The recombination event can take place in *E. coli* (Luckow *et al.*, 1993), yeast (Patel *et al.*, 1992), and *in vitro* (Ernst *et al.*, 1994). Subsequently, only modified recombinant baculoviral DNA is introduced into insect cell culture. An advantage of these technologies are that there is no need to further isolate a single recombinant virus by plaque assay since recombinant viral DNA is characterized prior to its introduction into cells.

Currently a variety of transfer vectors are available for baculovirus expression. These vectors have largely satisfied the need to ensure high-level synthesis of recombinant proteins (Jones and Morikawa, 1996). Transfer vectors to be used in *E. coli* contains an *E. coli* origin of replication, an ampicillin resistance marker, promoter (polyhedrin, p10, basic protein or pg64), and a cloning region downstream from the promoter in which to insert foreign genes. The purification of proteins from a recombinant source may also be a limiting factor in obtaining sufficient protein for desired structure and function studies. Thus, the inclusion of tags to enable purification of recombinant forms of protein has gained wide spread use. Some of these tags include glutathione S-transferase (GST), histidine amino acids, and cellulose binding domain (CBD_{cenD}).

1.6.2.2 Yeast expression systems

Yeast expression systems combine the ease, simplicity and low cost of bacterial expression systems with the authenticity of the far more expensive and less convenient animal tissue culture systems (Sudbery, 1996). The first yeast to be employed for the production of recombinant proteins was *S. cerevisiae.* Despite successes achieved with this organism, *S. cerevisiae* is often limited as an expression system by low yields. The methylotrophic yeast, *Pichia pastoris*, retain the advantages of *S. cerevisiae* but provides a reliable means of achieving greatly elevated expression yields.

In the past few years the *Pichia pastoris* system has gained wide interest since this system displays an extremely high yield of intracellular proteins. Percentage yields of between 5 – 40 % of total cell proteins are achieved routinely (Romanos, 1995). A further advantage of this system is the very efficient secretion of proteins into an almost protein-free extracellular medium. However, reports noting low yields or failure of expression in yeast are on the increase. These reports also indicate the proteolysis of

secreted polypeptides and insufficient secretion of complex foreign proteins since heterologous secretion of proteins is more demanding than intracellular expression. Expression of some genes does not yield any detectable protein, often because yeast transcriptional terminators result in truncated mRNA.

1.7 OBJECTIVES OF THE STUDY

Strains of *Clostridium botulinum* produce neurotoxins which cause fatal intoxication in both humans and animals. There are seven toxigenic types of *C. botulinum* (A to G) based on antigenically distinct toxins produced by different strains of the organism. Botulism of cattle, caused by types C and D neurotoxins, has a severe economic impact on cattle farming in South Africa and neighbouring countries. The currently available commercial vaccine relies on problematic anaerobic cultivation of the *Clostridium* bacterium followed by isolation, purification and inactivation of the toxin by treatment with formalin. Apart from the fastidious growth requirements of this organism, it has been reported that the production of toxin by these cells declines rapidly and eventually ceases, following laboratory passaging of the bacterial cultures. Thus, there exists a great need for a safe, efficacious and inexpensive vaccine against botulism. It is our principle aim to clone the type C and D botulinum neurotoxin genes, mutate the toxin genes, and to use molecular expression technologies for the expression of a possible recombinant vaccine candidate.

Specific objectives

i) To obtain full-length cloned copies of the *Clostridium botulinum* type C and D neurotoxin genes by making use of PCR and cloning methods.

ii) To obtain mutant versions of the full-length type C and D by making use of PCR-based site-directed mutagenesis.

iii) To express the wild-type and mutant neurotoxin genes in the BAC-TO-BAC™ baculovirus expression system.

CHAPTER 2

CLONING AND SITE-DIRECTED MUTAGENESIS OF NEUROTOXIN GENES OF *Clostridium botulinum* TYPES C AND D

2.1 INTRODUCTION

A century has passed since the neurotoxins produced by clostridia were identified as the causative agent of the syndrome called botulism. Currently, several *Clostridium* species are implicated in the production of seven serologically distinct botulinum neurotoxins (BoNT A to G). These botulinum neurotoxins are among the most toxic substances known to man (LD_{50} between 0.1 ng and 1.0 ng/kg)(Montecucco and Schiavo, 1994) and affect humans and animals alike. Ingestion of these neurotoxins result in partial or complete paralysis of the muscles of locomotion, mastication and deglutition (Kriek and Odendaal, 1994). In Southern Africa, Australia and South America serious outbreaks of botulism in livestock are commonly associated with the type C and D neurotoxins. These neurotoxins are a complex group of toxins that display differences in molecular structure and antigenicity within the same type (Moriishi *et al.*, 1989).

The understanding of the type C and D botulinum neurotoxin structure was facilitated with the realization that the genes for these toxins are bacteriophage carried. Vinet *et al.* (1968) first observed large phages with an octahedral head, sheath and tail in type C cultures. Since then a direct relationship has been established between the type C and D bacteriophages and toxicogenicity (Eklund *et al.*, 1971). Furthermore, it has been proposed that the phage exists in an extrachromosomal state (pseudolysogeny), somewhere between virulence and lysogeny (Minton, 1994). The above mentioned knowledge facilitated the reconstruction of the BoNT/C and BoNT/D genes from phage DNA libraries with the help of PCR amplification (Kimura *et al.*, 1990; Moriishi *et al.*, 1996a,b). Comparative analysis of the nucleotide sequences and derived amino acid sequences, obtained from these constructs, indicated that the type C and D neurotoxins are composed in a mosaic-like fashion and four isoforms sharing homology to different regions can be distinguished (Moriishi *et al.*, 1996a,b).

i 15785919
b15237745x

Investigation of the molecular structure and function of botulinum neurotoxins proved to be of great importance in the development of recombinant vaccines against botulism. Studies have shown that the botulinum neurotoxins are metallo-proteases that enter nerve cells and inhibit neurotransmitter release via zinc-dependent cleavage of protein components of the neuro-exocytosis apparatus (reviewed by Montecucco and Schiavo, 1994). The zinc-binding motif, which corresponds to typical zinc-endopeptidases (His-Glu-aa-aa-His), are located in the center of the light chain sub-unit. Kiyatkin *et al.* (1997) reported that site-directed mutagenesis of the zinc-binding histidine motif of BoNT/C resulted in the inactivation of this botulinum neurotoxin. The recombinant protein lacked toxicity *in vivo*, failed to block neuromuscular transmission *in vitro* and had no ability to cleave syntaxin in a broken-cell assay.

The aim of this study was to obtain full-length copies of the neurotoxin genes by employing polymerase chain reaction (PCR) amplification from bacteriophage DNA isolated from *Clostridium botulinum* type C (Stockholm) and D (South Africa) cultures. Using PCR-based site-directed mutagenesis procedures three amino acid mutations were introduced into the zinc-binding motif of the respective neurotoxins. The wild-type and mutant genes will subsequently be expressed in the BAC-TO-BAC™ baculovirus expression system for the development of a recombinant vaccine candidate.

2.2 MATERIALS AND METHODS

2.2.1 Materials

Clostridium botulinum type C_{ST} (Stockholm) and D_{SA} (South Africa) were used in this study. These *Clostridium* cultures, obtained from Design Biologix cc, were grown under anaerobic conditions in meat broth.

Primers used in PCR, sequencing and mutagenesis procedures were custom-synthesized by GIBCO BRL Life Technologies (Table 2.1). Commercially available primers were used for the sequencing of the 5' and 3' termini of the BoNT genes.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Table 2.1 Sequences of PCR primers used for amplification of BoNT genes

Primer	Polarity	Oligonucleotide sequence*
§botC-F	Sense	5'-GCGGAGCTCATGACAATAACAATTAACAAC-3'
‡botC-R	Anti-sense	5'-GCGGTCGACTTATTCACTTACAGGTAC-3'
§botD-F	Sense	5'-GCGGAGCTCATGACATTGCCAGTAAAAG-3'
†Cmut-F	Sense	5'-TTT**ggTacc**CTTAATAATGCAATG**c**ATAATTTAT-3'
†Cmut-R	Anti-sense	5'-TTT**ggTacc**CATTAAAATTAGTATTGGATC-3'
†Dmut-F	Sense	5'-TTT**ggTacc**TTAACAAATTCTTTG**c**ATCAATTG-3'
†Dmut-R	Anti-sense	5'-TTT**ggTacc**CATTAAAGCTATTACTGGATC-3'

§ Underlined are the unique *SacI* restriction enzyme sites included in design of primers.

‡ Underlined is the unique *SalI* restriction enzyme site included in design of primer.

† Underlined are the unique *KpnI* restriction enzyme sites included in design of primers.

* Mispaired bases are shown in bold lower case letters.

Table 2.2 Sequences of primers used for sequencing of full-length BoNT genes

Primer	Sequence	Source
-21M13	5'-TGTAAAACGACGGCCAGT-3'	Perkin-Elmer Biosystems
PUC Rev	5'-TCACACAGGAAACAGCTATGAC-3'	Promega
BNTseqC4	5'-CTAGACAAGGTAACCAACTGGGT-3'	Gibco BRL
BNTseqC13	5'-TCTACCATCTATTGCTTTATGAC-3'	Gibco BRL
BNTseqD4	5'-CGTCATACTACTAATATTGCAG-3'	Gibco BRL
BNTseqD13	5'-AATGCTATCTTTATCAGCTACAT-3'	Gibco BRL
BNTseq18	5'-GCTGATACATCTGATATTTTATC-3'	Gibco BRL
BNTseq23	5'-ATTTAATTCATCAATTACTTTAGG-3'	Gibco BRL
BNTseq28	5'-ATATCCAGGTAAATTACTTACCC-3'	Gibco BRL
BNTseq32	5'-TATCCTTCATTGAAGTCATTATTA-3'	Gibco BRL

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

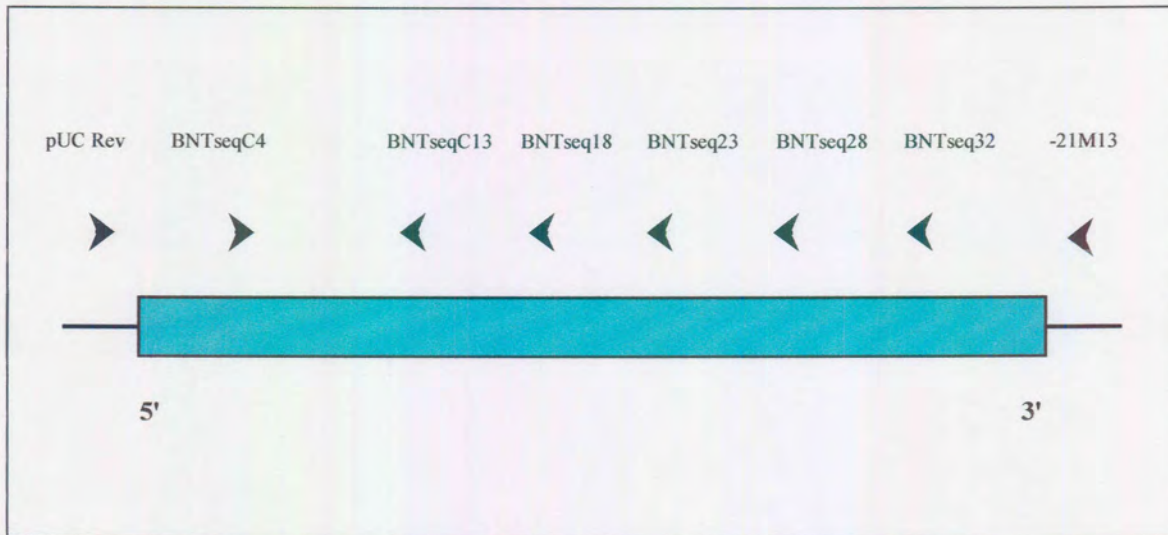

Fig. 2.1 Schematic presentation of specific primer annealing positions and direction of sequencing (</>), in the BoNT/C_{ST} gene and adjacent pGEM®-T Easy vector.

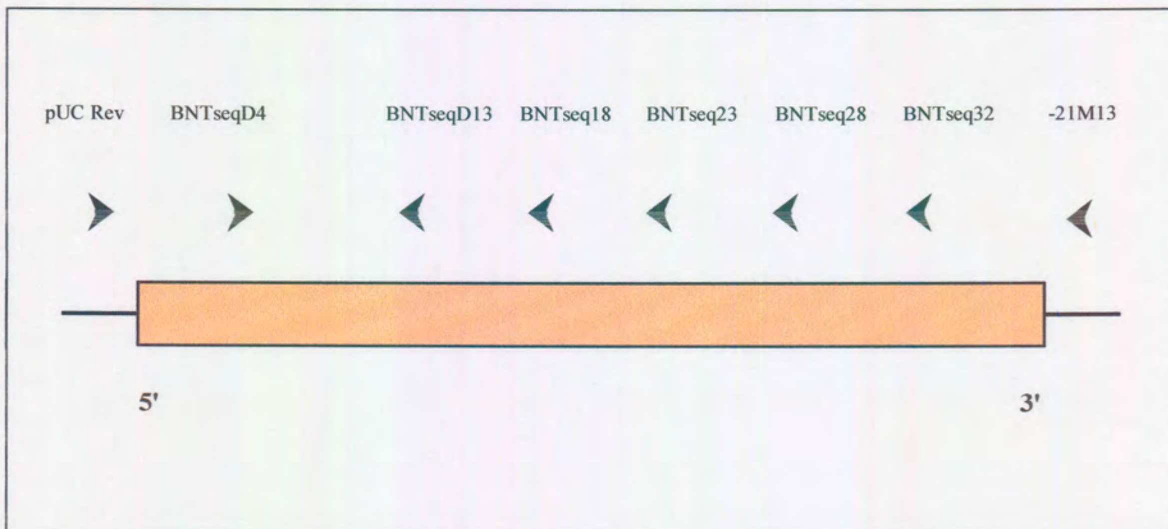

Fig. 2.2 Schematic presentation of specific primer annealing positions and direction of sequencing (</>), in the BoNT/D_{SA} gene and adjacent pGEM®-T Easy vector.

2.2.2 Transmission Electron Microscopy (TEM) of bacteriophages

Transmission electron microscopy (TEM) was used to confirm the presence of the bacteriophages within the *C. botulinum* cultures. Samples were fixed on copper grids and stained with 1% (w/v) potassium phosphate tungstic acid (KPTA) for 1 min after which they were viewed at 60 kV on a Phillips 3000 Transmission Electron Microscope at the Unit for Electron Microscopy, University of Pretoria.

2.2.3 Isolation of bacteriophage DNA

Bacteriophage DNA was isolated by phenol/chloroform extractions of the bacterial cultures. An equal volume of phenol was added to *C. botulinum* cultures in meat broth. The samples were centrifuged in a Hettich centrifuge at 13 000 rpm for 5 min after which an equal volume of phenol:chloroform (1:1) was added to the aqueous phase. Following centrifugation at 13 000 rpm for 5 min, the upper aqueous phase was extracted twice with an equal volume of chloroform. The DNA was subsequently precipitated from the supernatant at –70°C for 2 h following the addition of 2 volumes 96% ethanol and sodium acetate (pH 7.0) to a final concentration of 0.3 M. Precipitated DNA was pelleted by centrifugation at 13 000 rpm for 15 min, washed twice with 70% ethanol, dried under vacuum and resuspended in 1x TE buffer (10 mM Tris-HCl, pH 7.4; 1 mM EDTA, pH 8.0).

2.2.4 Agarose gel electrophoresis

Analysis of DNA was performed on 0.8% (w/v) SeaKem® LE agarose (FMC BioProducts) gels. The agarose gels were electrophoresed at 100 V in 1x TAE buffer (40 mM Tris-HCl; 2 mM EDTA; 20 mM acetic acid; pH 8.5). The gels were supplemented with EtBr (0.5 μg/ml) in order to allow visualization of the DNA on an UVItec transilluminator. Lamda (λ) DNA restricted with both *EcoRI* and *HindIII* (Promega) was used as molecular size marker to verify the size of DNA fragments. DNA concentrations were determined on a transilluminator according to the method described by Sambrook *et al.* (1989).

2.2.5 Amplification of botulinum neurotoxin genes

The BoNT/C_{ST} and BoNT/D_{SA} genes were amplified by polymerase chain reaction (PCR) amplification in a Perkin-Elmer 2400 thermal cycler (Perkin-Elmer, Cetus, Norwalk, CT, USA). The PCR mixture (50 µl) contained 5 µl 10x DNA polymerase buffer, 2 mM $MgCl_2$, 200 µM of each deoxynucleoside triphosphate (dNTP), 50 ng bacteriophage DNA, 10 pmol of the sense (botC-F or botD-F) and anti-sense primers (botC-R) (Table 2.1), 2.5 U of TaKaRa ExTaq™ DNA polymerase (TaKaRa Biomedicals), and ultra high quality (UHQ) water. The samples were subjected to 30 of the following cycles: denaturation at 94°C for 10 s, primer annealing at 49.5°C for 30 s, and elongation at 68°C for 3.5 min. In the final cycle an elongation step of 7 min was included in order to obtain full-length products. A negative control for which bacteriophage DNA was omitted was also included to test for contamination of solutions. The amplicons were analyzed on a 0.8% (w/v) agarose gel as described in section 2.2.4.

2.2.6 Purification of DNA fragments from agarose gels

The amplicons were purified from a 0.8% (w/v) agarose gel using a silica suspension as described by Boyle and Lew (1995). Briefly, both amplicons were excised from the agarose gel and mixed with 3 volumes of a 6 M NaI solution. The agarose was dissolved by incubation at 55°C for 5 min after which, 5 -10 µl of the silica suspension was added to each sample. The DNA was allowed to bind to the silica by incubation of the samples on ice for 30 min with intermittent mixing. The DNA-silica complex was pelleted by centrifugation and washed three times with NEW wash buffer (50 mM NaCl; 10 mM Tris-HCl, pH 7.5; 2.5 mM EDTA; 50% (v/v) ethanol). The DNA was eluted from the sillica in a final volume of 10 µl 1x TE buffer by incubation at 55°C for 3 min. The purified fragments were analyzed on a 0.8% (w/v) agarose gel as described previously.

2.2.7 Cloning of amplicons into the pGEM®-T Easy vector

2.2.7.1 Ligation of amplicons and plasmid vectors

The pGEM®-T Easy vector system (Promega) was used to clone the gel-purified BoNT amplicons. Ligation of the purified amplicons and the linear pGEM®-T Easy vector was

performed overnight (16 h) at 4°C in a total volume of 10 μl. The reaction mixture contained 1 μl of a 10x DNA ligase buffer (300 mM Tris-HCl, pH 7.8; 100 mM $MgCl_2$; 100 mM DTT; 10 mM ATP), 50 ng of pGEM®-T Easy vector, ca. 190 ng insert DNA, 3 U T4 DNA ligase (Promega, 3 U/μl), and UHQ water. The vector to insert ratio was typically 1:3.

2.2.7.2 Preparation of competent cells

Competent *E. coli* cells were prepared according to the $CaCl_2$ procedure as described by Sambrook *et al.* (1989). An overnight culture was prepared by inoculating 3 ml of LB broth (1% (w/v) tryptone; 0.5% (w/v) yeast extract; 1% (w/v) NaCl; pH 7.4) with a colony from a freshly streaked *E. coli* JM109 plate and shaking overnight at 37°C. One ml of the overnight culture was inoculated into 100 ml pre-heated LB broth (at 37°C) and grown to an OD_{550} value of 0.35 (mid-log phase). After incubation of the culture on ice for 30 min to inhibit further cell growth, cells from 35 ml were collected in a Corex tube by centrifugation at 4 000 rpm for 10 min at 4°C in a pre-cooled Sorvall HB4 rotor. The pellet was suspended in 10 ml ice-cold 50 mM $CaCl_2$, incubated on ice for 30 min, and pelleted as above. The pellet was finally suspended in 1 ml of the ice-cold 50 mM $CaCl_2$ solution and incubated on ice for at least 3 h before use in transformation experiments.

2.2.7.3 Transformation of competent cells

The competent *E. coli* JM109 cells were transformed after the addition of the ligation mixture using the heat shock method as described by Sambrook *et al.* (1989). The competent cells (200 μl) and the ligation reaction mixture (5 μl) were mixed in a glass transformation tube and maintained on ice for 30 min. The cells were then heat-shocked for 90 s at 42°C and chilled on ice for 2 min before pre-warmed LB broth (800 μl) was added. The transformation mixtures were incubated with agitation at 37°C for 1 h to allow the cells to recover and to express the ampicillin resistance gene. The transformed cells were selected by plating the cells onto LB-ampicillin plates (100 μg/ml) in the presence of 10 μl IPTG (100 mM stock solution) and 40 μl X-gal (2% (w/v) stock solution) to allow for blue/white colour selection. The plates were incubated overnight at 37°C and investigated for the presence of recombinant transformants with a Gal^- phenotype.

pUC18 plasmid DNA (10 ng) was included as a transformation control and to calculate the competency of the cells. A negative control consisting of competent cells only was also included to determine if any contamination occurred.

2.2.8 Plasmid DNA extraction

Plasmid DNA was isolated from selected colonies by the alkaline lysis method (Birnbiom and Doly, 1979) as described by Sambrook *et al.* (1989), with the following modifications. Colonies were picked from the agar plates with sterile toothpicks, inoculated into 5 ml of LB broth containing ampicillin (100 μg/ml), and then incubated overnight at 37°C with agitation. After incubation, cells from 2 ml of each culture was collected in Eppendorf tubes by centrifugation at 13 000 rpm for 1 min. The supernatant was discarded and the bacterial pellet suspended in 100 μl of Solution I (50 mM glucose; 25 mM Tris-HCl, pH 8.0; 10 mM EDTA, pH 8.0) before incubation on ice for 5 min. Two hundred microliters (200 μl) of Solution II (1% (w/v) SDS; 0.2 N NaOH) was then added and the tubes were placed on ice for 5 min. The tubes were gently inverted to avoid shearing of the DNA. Following the addition of 150 μl of Solution III (3 M sodium acetate, pH 4.8), the tubes were incubated on ice for 15 min. The genomic DNA, high molecular weight RNA, and cellular proteins were removed by centrifugation at 13 000 rpm for 15 min. The plasmid DNA was precipitated from the supernatant by the addition of 2 volumes of 96% ethanol at -20°C for 1h. Precipitated plasmid DNA was collected by centrifugation at 13 000 rpm for 15 min and washed twice with 70% ethanol to remove residual salts from the DNA. After drying the pellet under vacuum for 5 min to remove the excess ethanol, the pellets were resuspended in 1x TE buffer. The plasmid DNA were analyzed on a 0.8% (w/v) agarose gel.

2.2.9 Restriction enzyme digestion

Recombinant plasmid DNA was digested with *EcoRI* (Promega) and *SalI* (Roche Diagnostics) to determine the presence and size of DNA inserts. The reaction mixtures (15 – 20 μl) containing plasmid DNA (ca. 1 μg), 5 U of enzyme and the appropriate restriction buffer were incubated at 37ºC for 45 min. Resultant digested products were analyzed on a 0.8% (w/v) agarose gel as described in section 2.2.4.

2.2.10 Construction and cloning of mutant botulinum neurotoxin genes

2.2.10.1 Inverse PCR

Botulinum mutant neurotoxins, lacking 2 histidine and 1 glutamate residues in the zinc-binding region (His-Glu-aa-aa-His), were constructed by inverse PCR as described by Imai *et al.* (1991). This method utilizes two mutant oligonucleotide primers designed to amplify the target DNA (cloning vector as well as the target BoNT gene) in an inverted tail-to-tail direction, subsequently generating the desired mutation. A unique *KpnI* site was included in the design of the primers to generate substitutions of specific amino acids in the zinc-binding domain of the type C and D neurotoxins (Table 2.1). The Inverse PCR reactions were performed in a Perkin-Elmer 2400 thermal cycler. The PCR mixture (50 μl) contained 5 μl of 10x DNA polymerase buffer, 200 μM of each dNTP, 10 ng plasmid DNA, 10 ρmol of the sense and anti-sense mutant primers (Table 2.1), 2.5 U of TaKaRa ExTaq™ DNA polymerase (TaKaRa Biomedicals), and ultra high quality (UHQ) water. The samples were subjected to 30 of the following cycles: denaturation at 94°C for 10 s, primer annealing at 49.5°C for 30 s, and elongation at 68°C for 3.5 min. In the final cycle an elongation step of 7 min was included in order to obtain full-length products. A negative control from which plasmid DNA was omitted was included to test for contamination of solutions. The amplicons were then analyzed on a 0.8% (w/v) agarose gel as described in section 2.2.4.

2.2.10.2 Construction of clones containing mutated neurotoxin genes

Gel-purified amplicons were digested with *KpnI* (Roche Diagnostics) to generate compatible 5' and 3' ends for selfligation. Following agarose gel electrophoresis of the restricted DNA, bands were excised and purified with a silica suspension as described previously (section 2.2.6). The respective purified amplicons were self-ligated overnight at 15°C in a total volume of 10 μl. Ligation reaction mixtures were transformed into competent *E. coli* JM109 cells as described previously (section 2.2.7.3).

Recombinant plasmid DNA containing the mutated genes was purified by alkaline lysis method (section 2.2.8). The resultant mutant plasmids, designated pGEMut-C_{ST} and pGEMut-D_{SA}, were first characterized by *KpnI* restriction digestion and then by automated sequencing.

2.2.11 Automated sequencing of selected clones

2.2.11.1 Purification of plasmid DNA

Plasmid DNA for sequencing was purified as follows. Colonies were inoculated into 5 ml of LB broth containing ampicillin (100 μg/ml), and then incubated overnight at 37°C with agitation. After incubation, cells from 3 ml of each culture was collected in Eppendorf tubes by centrifugation at 13 000 rpm for 2 min. The supernatant was discarded and the bacterial pellet suspended in 400 μl of Solution I (50 mM glucose; 25 mM Tris-HCl, pH 8.0; 10 mM EDTA, pH 8.0; 100 ng RNAse) before incubation at room temperature for 10 min. Four hundred microliters (400 μl) of Solution II (1% (w/v) SDS; 0.2 N NaOH) was then added and the tubes were placed on ice for 10 min. The tubes were gently inverted to avoid shearing of the DNA. Following the addition of 300 μl of 7.5 M ammonium acetate (pH 7.6), the tubes were incubated on ice for 10 min before centrifugation at 13 000 rpm for 10 min. The plasmid DNA was precipitated from 1 ml of supernatant by the addition of 650 μl of isopropanol at room temperature for 10 min. Precipitated plasmid DNA was collected by centrifugation at 13 000 rpm for 10 min and the supernatant discarded before the addition of 100 μl of 2 M ammonium acetate (pH 7.4). The tubes were incubated on ice for 10 min and centrifuged at 13 000 rpm for 10 min. One hundred and ten microliters (110 μl) of isopropanol was added to the supernatant and the tubes were placed at room temperature for 10 min before centrifugation at 13 000 rpm for 10 min. The DNA pellet was washed twice with 70% ethanol to remove residual salts from the DNA. After drying the pellet under vacuum for 10 min to remove the excess ethanol, the pellets were resuspended in 1x TE buffer. The plasmid DNA were analyzed on a 0.8% (w/v) agarose gel as described in section 2.2.4.

2.2.11.2 Automated sequencing

The integrity of the full-length BoNT encoding genes and mutated regions were verified by automated sequencing of the clones using the ABI PRISM™ BigDye Terminator Cycle sequencing kit (Perkin-Elmer Biosystems). The sequencing reactions contained terminator Ready Reaction mix (A-dye terminator labeled with dichloro[R6G], C-dye terminator labeled with dichloro[ROX], G-dye terminator labeled with dichloro[R110], T-dye terminator labeled with dichloro[TAMRA]; deoxynucleoside triphosphates; AmpliTaq DNA polymerase FS; $MgCl_2$; Tris-HCl buffer, pH 9.0), 3.2 ρmol sequencing primer (Table

2.2), ca. 500 ng plasmid DNA, and UHQ water. Cycle sequencing reactions were performed in the Perkin-Elmer GeneAmp 2400 thermal cycler with 25 of the following cycles: denaturation at 96°C for 10 s, primer annealing at 50°C for 5 s, and elongation at 60°C for 4 min. Extension products were purified by centrifugation through a Sephadex G-50 (Fine) (Pharmacia) column at 3 000 rpm for 2 min. Samples were dried under vacuum for 30 min and stored at –20°C. Sequencing loading dye (5:1)(deionized formamide: 25 mM EDTA, pH 8.0 with blue dextran (50 mg/ml)) was added prior to electrophoresis of the samples on a Perkin-Elmer ABI 377 cycle sequencer.

2.3 RESULTS

2.3.1 Construction of clones containing the full-length copies of the C and D neurotoxin genes

The relationship between toxicogenicity and bacteriophage presence in type C and D cultures has long been established (Inoue and Iida, 1968; Eklund *et al.*, 1971) and it has been confirmed that type C and D phages carry the structural genes for the respective neurotoxins (Fujii *et al.,* 1988). In this study, the presence of bacteriophages in the *Clostridium* cultures was verified to ensure that the target DNA for PCR amplification of the full-length genes was available. During experimentation, new type C cultures had to be obtained after bacteriophage loss was demonstrated with TEM. Thereafter, bacteriophages with an octahedral head, sheath and tail (Fig. 2.3(A) and (B)) were observed in both cultures. The observed morphology of the phages corresponds to the type C and D bacteriophages reported by Inoue and Iida (1968) and Dolman and Chang (1972).

Phage DNA was extracted directly from the *C. botulinum* cultures by phenol-chloroform extractions and ethanol precipitation. Electrophoretic analysis on a 0.8% (w/v) agarose gel indicated two DNA bands for each phage DNA preparation. One band was localized in the well of the gel and the other migrated to a position above the 21 kb band of the DNA marker (results not shown). This corresponds to observations made by Fujii *et al.* (1988), which showed that both linear and open-circular DNAs are present in each phage DNA preparation. The type C and D bacteriophage DNA is estimated to be approximately 110 to 150 kb in length (Fujii *et al.*, 1988).

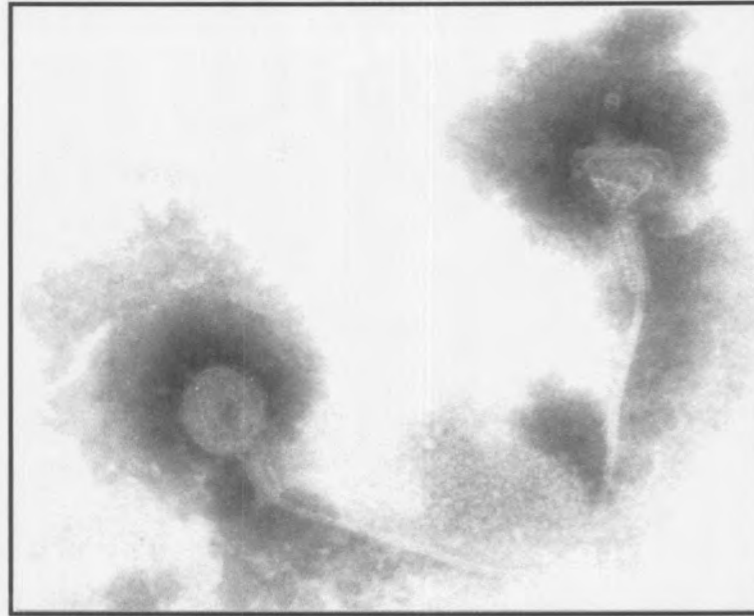

Fig. 2.3 (A) Transmission electron micrographs of bacteriophages from *Clostridium botulinum* type C_{ST}(Stockholm) cultures stained with KPTA.

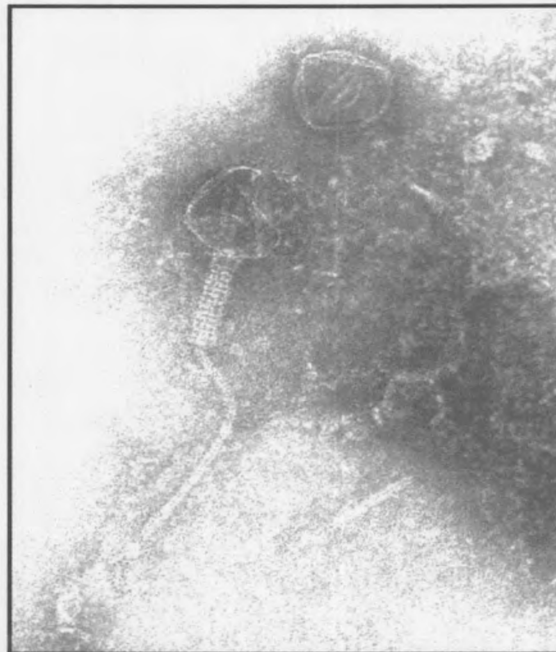

Fig. 2.3 (B) Transmission electron micrographs of bacteriophages from *Clostridium botulinum* type D_{SA}(South Africa) cultures stained with KPTA.

In order to obtain full-length copies of the respective neurotoxin genes, primers were designed based on published sequences of the type C and D botulinum neurotoxin genes (Binz *et al.*, 1990; Hauser *et al.*, 1990; Kimura *et al.*, 1990; Moriishi *et al.*, 1996a). The type C and D botulinum neurotoxin genes are composed in a mosaic-like fashion and four isoforms sharing homology to different regions can be distinguished. The BoNT/C_{ST} and BoNT/D_{SA} genes have different 5' terminal ends, but share a homologous 3' terminal end. Primers were subsequently designed based on the sequences of these regions. Each of the primers were extended at their 5' ends by an additional 9 nucleotides of which specified either *SalI* or *SacI* restriction sites (Table 2.1) with an additional 3 nucleotide clamping sequence to facilitate more efficient restriction enzyme digestion. Restriction enzyme sites (*SacI* and *SalI*) were incorporated to facilitate cloning procedures.

Using the isolated bacteriophage DNA preparations as template, PCR amplification was carried out. Primer sets designed according to published sequences of the BoNT/C and BoNT/D genes were used in conjunction with a polymerase enzyme specifically developed for long and accurate (LA) PCR technology. Initial attempts in amplifying the genes using Taq DNA polymerase (Promega) and AmpliTaq Gold (Perkin-Elmer Biosystems) proved to be unsuccessful. However, the use of the long-range TaKaRa Ex Taq™ (TaKaRa Biomedical) polymerase resulted in amplification. A single discreet band of 3.8 kb was observed following agarose gel electrophoresis (Fig. 2.4). No amplification products were observed in the negative control in which template DNA was omitted.

The gel-purified PCR amplicons were cloned into the pGEM®-T Easy vector (Promega) (Fig. 2.5) that contain 3' thymidine overhangs in order to facilitate the cloning of PCR products since a single non-template specific deoxyadenosine is added to the 3'-ends of the amplified fragments during PCR (Clark, 1988). In addition, the pGEM®-T Easy vector also contains two selectable markers, amp^R and *lacZ'*. These were used for the selection of recombinant transformants. Successful cloning of a gene into this vector disrupts the coding sequence of the α-peptide region of the enzyme β-galactosidase. This allows for the identification of recombinant clones by colorimetric screening on X-gal containing indicator plates. The isolated plasmid DNA was characterized by agarose gel electrophoresis and restriction enzyme analysis.

Plasmid DNA migrating slower that the parental pGEM®-T Easy vector was selected and digested with *SalI* (Roche Diagnostics) to determine whether the plasmid DNA contained cloned inserts. *SalI* digestion of pGEM-D_{SA} resulted in two fragments of ca. 3.8 kb and 3.0 kb, respectively (Fig. 2.6). This corresponds with the expected size of the DNA insert and parental linearized vector. Similarly, *SalI* digestion of pGEM-C_{ST} also resulted in two fragments with the above sizes. Recombinant plasmid DNA containing either the C_{ST} and D_{SA} genes were selected and used for the further DNA manipulations. The clones were designated pGEM-C_{ST} and pGEM-D_{SA}, respectively.

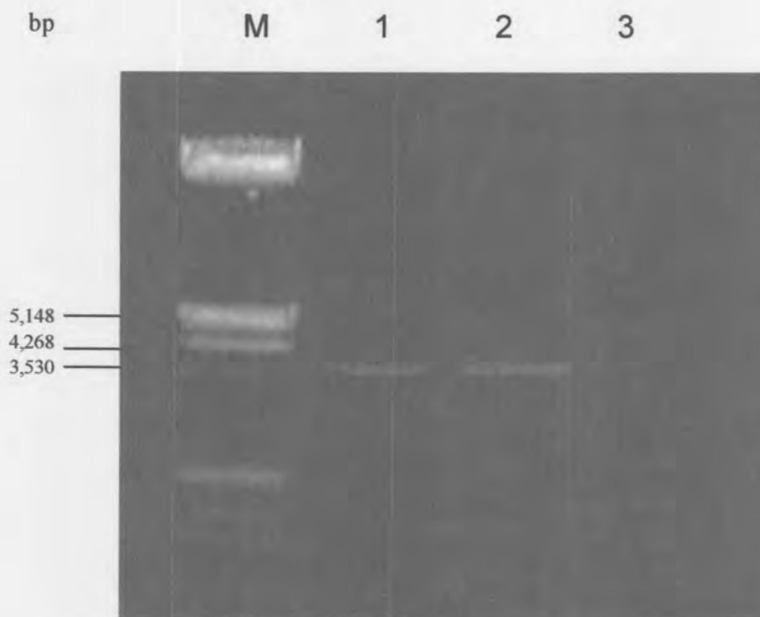

Fig. 2.4 Agarose gel analysis of the 3.8 kb fragment obtained by PCR amplification using bacteriophage DNA as template. Lane: M, λ DNA/*EcoRI* + *HindIII* Marker; 1, BoNT/C_{ST} amplicon; 2, BoNT/D_{SA} amplicon; 3, Negative control.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

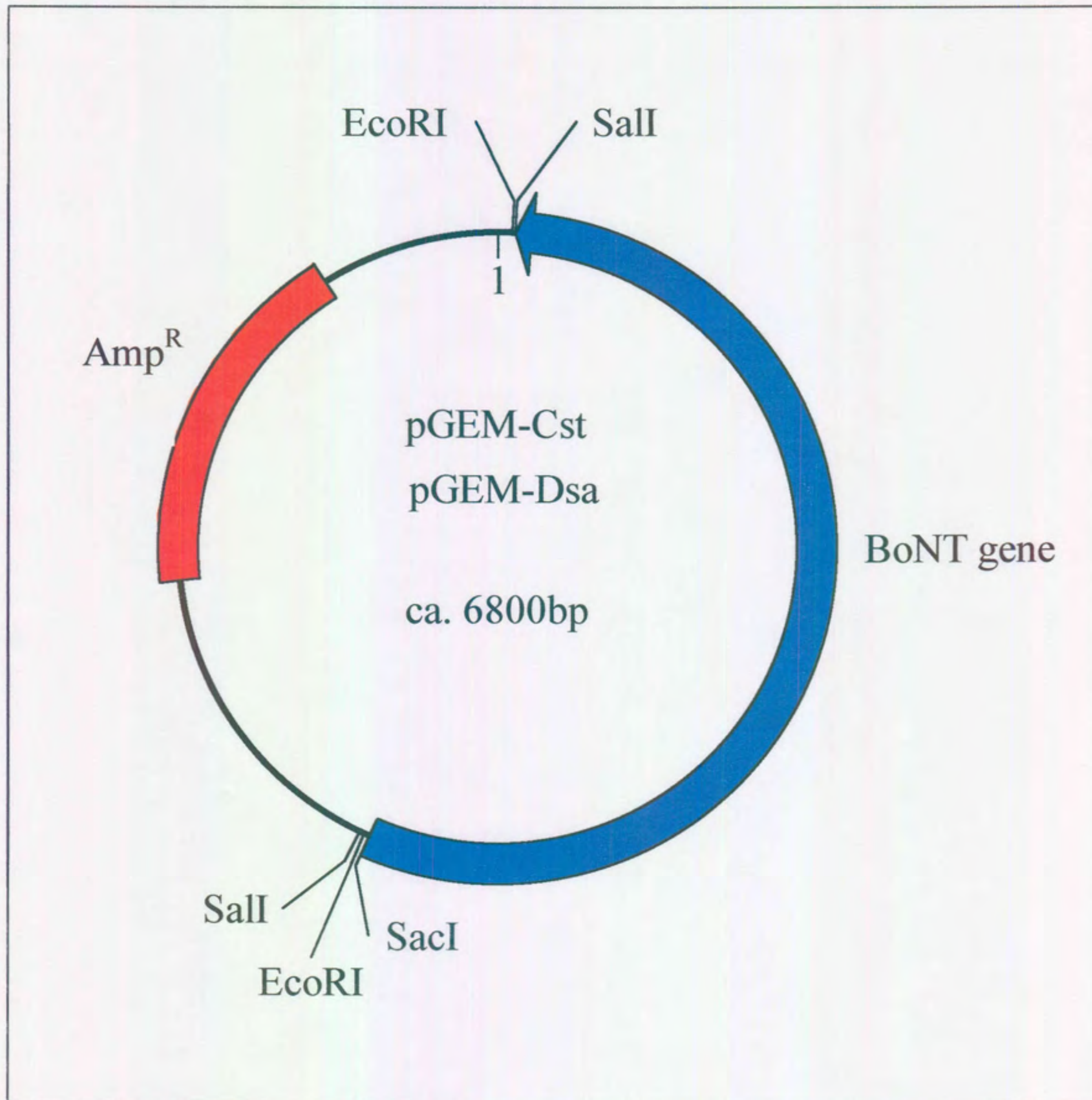

Fig. 2.5 Partial restriction enzyme map of recombinant pGEM®-T Easy vector containing either the BoNT/C_{ST} or the BoNT/D_{SA} gene.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

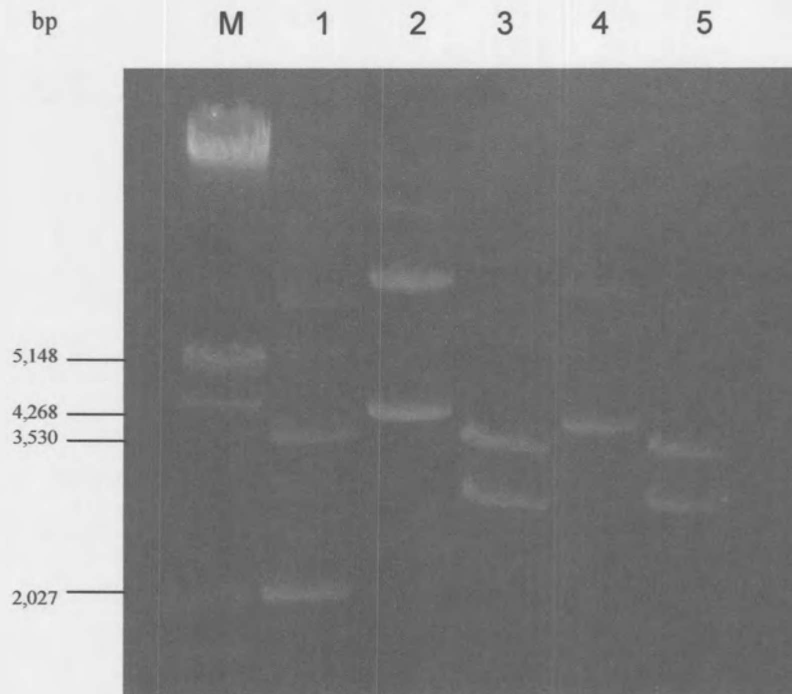

Fig. 2.6 Agarose gel electrophoretic analysis of the recombinant plasmids, pGEM-C_{ST} and pGEM-D_{SA}, constructed by cloning the respective BoNT amplicons into linearized pGEM®-T Easy vector. Lanes: M, λ DNA/*EcoRI* + *HindIII* Marker ; 1, Uncut non-recombinant pGEM®-T Easy vector; 2, Uncut pGEM-C_{ST} vector; 3, *SalI*-digested pGEM-C_{ST} vector; 4, Uncut pGEM-D_{SA} vector; 5, *SalI*-digested pGEM-D_{SA} vector.

2.3.2 Sequencing of full-length botulinum neurotoxin genes

The integrity of the full-length botulinum neurotoxin genes was determined by automated sequencing. The strategy used for sequencing of the full-length genes involved the design of intemal primers based on published sequences of the type C and D botulinum neurotoxin genes (Binz *et al.*, 1990; Hauser *et al.*, 1990; Kimura *et al.*, 1990; Moriishi *et al.*, 1996(b)). The nucleotide sequences of BoNT/C_{ST} and BoNT/D_{SA} genes were aligned and a set of four primers (BNTseq18, BNTseq23, BNTseq28, and BNTseq32) was subsequently designed to homologous regions. Furthermore, a set of two primers each (BNTseqC4 and BNTseqC13 / BNTseqD4 and BNTseqD13) was designed for the respective genes.

The obtained nucleotide sequences for the respective clones are shown in Fig. 2.7 and Fig. 2.8. Analysis of the nucleotide sequence indicated the presence of nucleotides specifying *SacI/SalI* restriction enzyme sites at the 5' and 3' termini of the genes. The nucleotide sequences were compared with the published sequences of the respective genes (Kimura *et al.*, 1990 (GI: 217780); Moriishi *et al.*, 1996a (GI: 1374775)). Sequence homologies of 99.8% (BoNT/C_{ST}) and 99.7% (BoNT/D_{SA}) were observed between the sequence obtained for the cloned genes and published sequences.

The deduced amino acid sequence of BoNT/C_{ST} displayed 5 amino acid dissimilarities with the amino acid sequence of the published sequence (Table 2.3). The charge of three of the dissimilar amino acids differed from the amino acids in the published sequence, while the other two remained the same. The deduced amino acid sequence of BoNT/D_{SA} displayed 7 amino acid dissimilarities with the amino acid sequence of the published sequence (Table 2.4). The charge of all of the dissimilar amino acids was the same as the amino acids of the published sequence. Comparison of the amino acid sequences of the four isoforms (BoNT/C_1, BoNT/C_{6813}, BoNT/D and BoNT/D_{SA}) (Moriishi *et al.*, 1996a,b) showed that several of the dissimilar amino acids contained in the cloned genes are similar to amino acids in the other isoforms. The Ala^{657} of the cloned BoNT/C_{ST} gene is similar to that of BoNT/D and BoNT/D_{SA}. The Val^{362} of the cloned BoNT/D_{SA} gene is similar to that of BoNT/D. Furthermore, Val^{917} and Val^{1283} are similar to BoNT/C_{ST}. The substitution of Ala^{1283} with Val^{1283} was as a result of the botC-R primer

used for amplification of the BoNT/D_{SA} gene. This primer was designed according to the sequence of BoNT/C_{ST}.

Table 2.3 Amino acid differences between cloned BoNT/C_{ST} and published sequence (Kimura *et al.,* 1990).

Kimura *et al.* (1990)		**Cloned C_{ST} gene**	
Amino acid	**Charge**	**Amino acid**	**Charge**
Pro^{2}	Non-polar	Thr^{2}	Polar
Pro^{196}	Non-polar	Leu^{196}	Non-polar
Thr^{657}	Polar	Ala^{657}	Non-polar
Ile^{686}	Non-polar	Val^{686}	Non-polar
Ser^{730}	Polar	Pro^{730}	Non-polar

Table 2.4 Amino acid differences between cloned BoNT/D_{SA} and published sequence (Moriishi *et al.,* 1996a).

Moriishi *et al.*, 1996a		**Cloned D_{SA} gene**	
Amino acid	**Charge**	**Amino acid**	**Charge**
Trp^{3}	Non-polar	Leu^{3}	Non-polar
Leu^{341}	Non-polar	Ala^{341}	Non-polar
Ile^{362}	Non-polar	Val^{362}	Non-polar
Ser^{755}	Polar	Gly^{755}	Polar
Ile^{917}	Non-polar	Val^{917}	Non-polar
Ile^{1042}	Non-polar	Val^{1042}	Non-polar
Ala^{1283}	Non-polar	Val^{1283}	Non-polar

```
                    10         20         30         40         50         60         70         80         90         100        110        120        130
1 C-ST (Ref) atgccaataa caattaacaa ctttaattat tcagatcctg ttgataataa aaatatttta tatttagata ctcatttaaa tacactagct aatgagcctg aaaaagcctt tcgcattaca ggaaatatat
2             M  P  I    T  I  N  N   F  N  Y    S  D  P    V  D  N  K   N  I  L    Y  L  D    T  H  L  N   T  L  A    N  E  P    E  K  A  F   R  I  T    G  N  I
3 pGEM-Cst   atgAcaataa caattaacaa ctttaattat tcagatcctg ttgataataa aaatatttta tatttagata ctcatttaaa tacactagct aatgagcctg aaaaagcctt tcgcattaca ggaaatatat
4             M  T  I    T  I  N  N   F  N  Y    S  D  P    V  D  N  K   N  I  L    Y  L  D    T  H  L  N   T  L  A    N  E  P    E  K  A  F   R  I  T    G  N  I
5            ---*------ ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                    140        150        160        170        180        190        200        210        220        230        240        250        260
1 C-ST (Ref) gggtaatacc tgatagattt tcaagaaatt ctaatccaaa tttaaataaa cctcctcgag ttacaagccc taaaagtggt tattatgatc ctaattattt gagtactgat tctgacaaag atccattttt
2            W  V  I  P   D  R  F    S  R  N    S  N  P  N   L  N  K    P  P  R    V  T  S  P   K  S  G    Y  Y  D    P  N  Y  L   S  T  D    S  D  K    D  P  F  L
3 pGEM-Cst   gggtaatacc tgatagattt tcaagaaatt ctaatccaaa tttaaataaa cctcctcgag ttacaagccc taaaagtggt tattatgatc ctaattattt gagtactgat tctgacaaag atccattttt
4            W  V  I  P   D  R  F    S  R  N    S  N  P  N   L  N  K    P  P  R    V  T  S  P   K  S  G    Y  Y  D    P  N  Y  L   S  T  D    S  D  K    D  P  F  L
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                    270        280        290        300        310        320        330        340        350        360        370        380        390
1 C-ST (Ref) aaaagaaatt ataaagttat ttaaaagaat taattctaga gaaataggag aagaattaat atatagactt tcgacagata taccctttcc tgggaataac aatactccaa ttaatacttt tgattttgat
2             K  E  I    I  K  L    F  K  R  I   N  S  R    E  I  G    E  E  L  I   Y  R  L    S  T  D    I  P  F  P   G  N  N    N  T  P    I  N  T  F   D  F  D
3 pGEM-Cst   aaaagaaatt ataaagttat ttaaaagaat taattctaga gaaataggag aagaattaat atatagactt tcgacagata taccctttcc tgggaataac aatactccaa ttaatacttt tgattttgat
4             K  E  I    I  K  L    F  K  R  I   N  S  R    E  I  G    E  E  L  I   Y  R  L    S  T  D    I  P  F  P   G  N  N    N  T  P    I  N  T  F   D  F  D
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                    400        410        420        430        440        450        460        470        480        490        500        510        520
1 C-ST (Ref) gtagatttta acagtgttga tgttaaaact agacaaggta acaactgggt taaaactggt agcataaatc ctagtgttat aataactgga cctagagaaa acattataga tccagaaact tctacgttta
2             V  D  F    N  S  V  D   V  K  T    R  Q  G    N  N  W  V   K  T  G    S  I  N    P  S  V  I   I  T  G    P  R  E    N  I  I  D   P  E  T    S  T  F
3 pGEM-Cst   gtagatttta acagtgttga tgttaaaact agacaaggta acaactgggt taaaactggt agcataaatc ctagtgttat aataactgga cctagagaaa acattataga tccagaaact tctacgttta
4             V  D  F    N  S  V  D   V  K  T    R  Q  G    N  N  W  V   K  T  G    S  I  N    P  S  V  I   I  T  G    P  R  E    N  I  I  D   P  E  T    S  T  F
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                    530        540        550        560        570        580        590        600        610        620        630        640        650
1 C-ST (Ref) aattaactaa caatactttt gcggcacaag aaggatttgg tgctttatca ataatttcaa tatcacctag atttatgcta acatatagta atgcaactaa tgatgtagga gagggtagat tttctaagtc
2            K  L  T  N   N  T  F    A  A  Q    E  G  F  G   A  L  S    I  I  S    I  S  P  R   F  M  L    T  Y  S    N  A  T  N   D  V  G    E  G  R    F  S  K  S
3 pGEM-Cst   aattaactaa caatactttt gcggcacaag aaggatttgg tgctttatca ataatttcaa tatcacTtag atttatgcta acatatagta atgcaactaa tgatgtagga gagggtagat tttctaagtc
4            K  L  T  N   N  T  F    A  A  Q    E  G  F  G   A  L  S    I  I  S    I  S  L  R   F  M  L    T  Y  S    N  A  T  N   D  V  G    E  G  R    F  S  K  S
5            ---------- ---------- ---------- ---------- ---------- ---------- ------*--- ---------- ---------- ---------- ---------- ---------- ----------

                    660        670        680        690        700        710        720        730        740        750        760        770        780
1 C-ST (Ref) tgaattttgc atggatccaa tactaatttt aatgcatgaa cttaatcatg caatgcataa tttatatgga atagctatac caaatgatca aacaatttca tctgtaacta gtaatatttt ttattctcaa
2             E  F  C    M  D  P    I  L  I  L   M  H  E    L  N  H    A  M  H  N   L  Y  G    I  A  I    P  N  D  Q   T  I  S    S  V  T    S  N  I  F   Y  S  Q
3 pGEM-Cst   tgaattttgc atggatccaa tactaatttt aatgcatgaa cttaatcatg caatgcataa tttatatgga atagctatac caaatgatca aacaatttca tctgtaacta gtaatatttt ttattctcaa
4             E  F  C    M  D  P    I  L  I  L   M  H  E    L  N  H    A  M  H  N   L  Y  G    I  A  I    P  N  D  Q   T  I  S    S  V  T    S  N  I  F   Y  S  Q
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                    790        800        810        820        830        840        850        860        870        880        890        900        910
1 C-ST (Ref) tataatgtga aattagagta tgcagaaata tatgcatttg gaggtccaac tatagacctt attcctaaaa gtgcaaggaa atattttgag gaaaaggcat tggattatta tagatctata gctaaaagac
2             Y  N  V    K  L  E  Y   A  E  I    Y  A  F    G  G  P  T   I  D  L    I  P  K    S  A  R  K   Y  F  E    E  K  A    L  D  Y  Y   R  S  I    A  K  R
3 pGEM-Cst   tataatgtga aattagagta tgcagaaata tatgcatttg gaggtccaac tatagacctt attcctaaaa gtgcaaggaa atattttgag gaaaaggcat tggattatta tagatctata gctaaaagac
4             Y  N  V    K  L  E  Y   A  E  I    Y  A  F    G  G  P  T   I  D  L    I  P  K    S  A  R  K   Y  F  E    E  K  A    L  D  Y  Y   R  S  I    A  K  R
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                    920        930        940        950        960        970        980        990        1000       1010       1020       1030       1040
1 C-ST (Ref) ttaatagtat aactactgca aatccttcaa gctttaataa atatataggg gaatataaac agaaacttat tagaaagtat agattcgtag tagaatcttc aggtgaagtt acagtaaatc gtaataagtt
2            L  N  S  I   T  T  A    N  P  S    S  F  N  K   Y  I  G    E  Y  K    Q  K  L  I   R  K  Y    R  F  V    V  E  S  S   G  E  V    T  V  N    R  N  K  F
3 pGEM-Cst   ttaatagtat aactactgca aatccttcaa gctttaataa atatataggg gaatataaac agaaacttat tagaaagtat agattcgtag tagaatcttc aggtgaagtt acagtaaatc gtaataagtt
4            L  N  S  I   T  T  A    N  P  S    S  F  N  K   Y  I  G    E  Y  K    Q  K  L  I   R  K  Y    R  F  V    V  E  S  S   G  E  V    T  V  N    R  N  K  F
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

```
             1050       1060       1070       1080       1090       1100       1110       1120       1130       1140       1150       1160       1170
1 C-ST (Ref) tgttgagtta tataatgaac ttacacaaat atttacagaa tttaactacg ctaaaatata taatgtacaa aataggaaaa tatatctttc aaatgtatat actccggtta cggcgaatat attagacgat
2             V  E  L   Y  N  E   L  T  Q  I   F  T  E   F  N  Y   A  K  I  Y   N  V  Q   N  R  K   I  Y  L  S   N  V  Y   T  P  V   T  A  N  I   L  D  D
3 pGEM-Cst    tgttgagtta tataatgaac ttacacaaat atttacagaa tttaactacg ctaaaatata taatgtacaa aataggaaaa tatatctttc aaatgtatat actccggtta cggcgaatat attagacgat
4             V  E  L   Y  N  E   L  T  Q  I   F  T  E   F  N  Y   A  K  I  Y   N  V  Q   N  R  K   I  Y  L  S   N  V  Y   T  P  V   T  A  N  I   L  D  D
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

             1180       1190       1200       1210       1220       1230       1240       1250       1260       1270       1280       1290       1300
1 C-ST (Ref) aatgtttatg atatacaaaa tggatttaat atacctaaaa gtaatttaaa tgtactattt atgggtcaaa atttatctcg aaatccagca ttaagaaaag tcaatcctga aaatatgctt tatttattta
2             N  V  Y   D  I  Q  N   G  F  N   I  P  K   S  N  L  N   V  L  F   M  G  Q   N  L  S  R   N  P  A   L  R  K   V  N  P  E   N  M  L   Y  L  F
3 pGEM-Cst    aatgtttatg atatacaaaa tggatttaat atacctaaaa gtaatttaaa tgtactattt atgggtcaaa atttatctcg aaatccagca ttaagaaaag tcaatcctga aaatatgctt tatttattta
4             N  V  Y   D  I  Q  N   G  F  N   I  P  K   S  N  L  N   V  L  F   M  G  Q   N  L  S  R   N  P  A   L  R  K   V  N  P  E   N  M  L   Y  L  F
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

             1310       1320       1330       1340       1350       1360       1370       1380       1390       1400       1410       1420       1430
1 C-ST (Ref) caaaatttg  tcataaagca atagatggta gatcattata taataaaaca ttagattgta gagagctttt agttaaaaat actgacttac cctttatagg tgatattagt gatgttaaaa ctgatatatt
2            T  K  F  C   H  K  A   I  D  G   R  S  L  Y   N  K  T   L  D  C   R  E  L  L   V  K  N   T  D  L   P  F  I  G   D  I  S   D  V  K   T  D  I  F
3 pGEM-Cst   caaaatttg  tcataaagca atagatggta gatcattata taataaaaca ttagattgta gagagctttt agttaaaaat actgacttac cctttatagg tgatattagt gatgttaaaa ctgatatatt
4            T  K  F  C   H  K  A   I  D  G   R  S  L  Y   N  K  T   L  D  C   R  E  L  L   V  K  N   T  D  L   P  F  I  G   D  I  S   D  V  K   T  D  I  F
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

             1440       1450       1460       1470       1480       1490       1500       1510       1520       1530       1540       1550       1560
1 C-ST (Ref) tttaagaaaa gatattaatg aagaaactga agttatatac tatccggaca atgtttcagt agatcaagtt attctcagta agaatacctc agaacatgga caactagatt tattataccc tagtattgac
2             L  R  K   D  I  N   E  E  T  E   V  I  Y   Y  P  D   N  V  S  V   D  Q  V   I  L  S   K  N  T  S   E  H  G   Q  L  D   L  L  Y  P   S  I  D
3 pGEM-Cst   tttaagaaaa gatattaatg aagaaactga agttatatac tatccggaca atgtttcagt agatcaagtt attctcagta agaatacctc agaacatgga caactagatt tattataccc tagtattgac
4             L  R  K   D  I  N   E  E  T  E   V  I  Y   Y  P  D   N  V  S  V   D  Q  V   I  L  S   K  N  T  S   E  H  G   Q  L  D   L  L  Y  P   S  I  D
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

             1570       1580       1590       1600       1610       1620       1630       1640       1650       1660       1670       1680       1690
1 C-ST (Ref) agtgagagtg aaatattacc aggggagaat caagtctttt atgataatag aactcaaaat gttgattatt tgaattctta ttattaccta gaatctcaaa aactaagtga taatgttgaa gattttactt
2             S  E  S   E  I  L  P   G  E  N   Q  V  F   Y  D  N  R   T  Q  N   V  D  Y   L  N  S  Y   Y  Y  L   E  S  Q   K  L  S  D   N  V  E   D  F  T
3 pGEM-Cst   agtgagagtg aaatattacc Ggggagaat  caagtctttt atgataatag aactcaaaat gttgattatt tgaattctta ttattaccta gaatctcaaa aactaagtga taatgttgaa gattttactt
4             S  E  S   E  I  L  P   G  E  N   Q  V  F   Y  D  N  R   T  Q  N   V  D  Y   L  N  S  Y   Y  Y  L   E  S  Q   K  L  S  D   N  V  E   D  F  T
5            ---------- ---------- *--------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

             1700       1710       1720       1730       1740       1750       1760       1770       1780       1790       1800       1810       1820
1 C-ST (Ref) ttacgagatc aattgaggag gctttggata atagtgcaaa agtatatact tactttccta cactagctaa taaagtaaat gcgggtgttc aaggtggttt atttttaatg tgggcaaatg atgtagttga
2            F  T  R  S   I  E  E   A  L  D   N  S  A  K   V  Y  T   Y  F  P   T  L  A  N   K  V  N   A  G  V   Q  G  G  L   F  L  M   W  A  N   D  V  V  E
3 pGEM-Cst   ttacgagatc aattgaggag gctttggata atagtgcaaa agtatatact tactttccta cactagctaa taaagtaaat gcgggtgttc aaggtggttt atttttaatg tgggcaaatg atgtagttga
4            F  T  R  S   I  E  E   A  L  D   N  S  A  K   V  Y  T   Y  F  P   T  L  A  N   K  V  N   A  G  V   Q  G  G  L   F  L  M   W  A  N   D  V  V  E
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

             1830       1840       1850       1860       1870       1880       1890       1900       1910       1920       1930       1940       1950
1 C-ST (Ref) agattttact acaaatattc taagaaaaga tacattagat aaaatatcag atgtatcagc tattattccc tatataggac ccgcattaaa tataagtaat tctgtaagaa gaggaaattt tactgaagca
2             D  F  T   T  N  I   L  R  K  D   T  L  D   K  I  S   D  V  S  A   I  I  P   Y  I  G   P  A  L  N   I  S  N   S  V  R   R  G  N  F   T  E  A
3 pGEM-Cst   agattttact acaaatattc taagaaaaga tacattagat aaaatatcag atgtatcagc tattattccc tatataggac ccgcattaaa tataagtaat tctgtaagaa gaggaaattt tactgaagca
4             D  F  T   T  N  I   L  R  K  D   T  L  D   K  I  S   D  V  S  A   I  I  P   Y  I  G   P  A  L  N   I  S  N   S  V  R   R  G  N  F   T  E  A
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

             1960       1970       1980       1990       2000       2010       2020       2030       2040       2050       2060       2070       2080
1 C-ST (Ref) tttgcagtta ctggtgtaac tattttatta gaagcatttc ctgaatttac aatacctgca cttggtgcat ttgtgattta tagtaaggtt caagaaagaa acgagattat taaaactata gataattgtt
2             F  A  V   T  G  V  T   I  L  L   E  A  F   P  E  F  T   I  P  A   L  G  A   F  V  I  Y   S  K  V   Q  E  R   N  E  I  I   K  T  I   D  N  C
3 pGEM-Cst   tttgcagtta ctggtgtaGc tattttatta gaagcatttc ctgaatttac aatacctgca cttggtgcat ttgtgattta tagtaaggtt caagaaagaa acgagGttat taaaactata gataattgtt
4             F  A  V   T  G  V  A   I  L  L   E  A  F   P  E  F  T   I  P  A   L  G  A   F  V  I  Y   S  K  V   Q  E  R   N  E  V  I   K  T  I   D  N  C
5            ---------- --------*- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----*---- ---------- ----------
```

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

```
                2090       2100       2110       2120       2130       2140       2150       2160       2170       2180       2190       2200       2210
1 C-ST (Ref) tagaacaaag gattaagaga tggaaagatt catatgaatg gatgatggga acgtggttat ccaggattat tactcaattt aataatataa gttatcaaat gtatgattct ttaaattatc aggcaggtgc
2            L  E  Q  R   I  K  R   W  K  D   S  Y  E  W   M  M  G   T  W  L   S  R  I  I   T  Q  F   N  N  I   S  Y  Q  M   Y  D  S   L  N  Y   Q  A  G  A
3 pGEM-Cst   tagaacaaag gattaagaga tggaaagatt catatgaatg gatgatggga acgtggttat ccaggattat tactcaattt aataatataa gttatcaaat gtatgatCct ttaaattatc aggcaggtgc
4            L  E  Q  R   I  K  R   W  K  D   S  Y  E  W   M  M  G   T  W  L   S  R  I  I   T  Q  F   N  N  I   S  Y  Q  M   Y  D  P   L  N  Y   Q  A  G  A
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -------*-- ---------- ----------

                2220       2230       2240       2250       2260       2270       2280       2290       2300       2310       2320       2330       2340
1 C-ST (Ref) aatcaaagct aaaatagatt tagaatataa aaaatattca ggaagtgata aagaaaatat aaaaagtcaa gttgaaaatt taaaaaatag tttagatgta aaaatttcgg aagcaatgaa taatataaat
2             I  K  A   K  I  D   L  E  Y  K   K  Y  S   G  S  D   K  E  N  I   K  S  Q   V  E  N   L  K  N  S   L  D  V   K  I  S   E  A  M  N   N  I  N
3 pGEM-Cst   aatcaaagct aaaatagatt tagaatataa aaaatattca ggaagtgata aagaaaatat aaaaagtcaa gttgaaaatt taaaaaatag tttagatgta aaaatCtcgg aagcaatgaa taatataaat
4             I  K  A   K  I  D   L  E  Y  K   K  Y  S   G  S  D   K  E  N  I   K  S  Q   V  E  N   L  K  N  S   L  D  V   K  I  S   E  A  M  N   N  I  N
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----*---- ---------- ----------

                2350       2360       2370       2380       2390       2400       2410       2420       2430       2440       2450       2460       2470
1 C-ST (Ref) aaatttatac gagaatgttc cgtaacatat ttatttaaaa atatgttacc taaagtaatt gatgaattaa atgagtttga tcgaaatact aaagcaaaat taattaatct tatagatagt cataatatta
2             K  F  I   R  E  C  S   V  T  Y   L  F  K   N  M  L  P   K  V  I   D  E  L   N  E  F  D   R  N  T   K  A  K   L  I  N  L   I  D  S   H  N  I
3 pGEM-Cst   aaatttatac gagaatgttc cgtaacatat ttatttaaaa atatgttacc taaagtaatt gatgaattaa atgagtttga tcgaaatact aaagcaaaat taattaatct tatagatagt cataatatta
4             K  F  I   R  E  C  S   V  T  Y   L  F  K   N  M  L  P   K  V  I   D  E  L   N  E  F  D   R  N  T   K  A  K   L  I  N  L   I  D  S   H  N  I
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                2480       2490       2500       2510       2520       2530       2540       2550       2560       2570       2580       2590       2600
1 C-ST (Ref) ttctagttgg tgaagtagat aaattaaaag caaaagtaaa taatagcttt caaaatacaa taccctttaa tattttttca tatactaata attctttatt aaaagatata attaatgaat atttcaataa
2            I  L  V  G   E  V  D   K  L  K   A  K  V  N   N  S  F   Q  N  T   I  P  F  N   I  F  S   Y  T  N   N  S  L  L   K  D  I   I  N  E   Y  F  N  N
3 pGEM-Cst   ttctagttgg tgaagtagat aaattaaaag caaaagtaaa taatagcttt caaaatacaa taccctttaa tattttttca tatactaata attctttatt aaaagatata attaatgaat atttcaataa
4            I  L  V  G   E  V  D   K  L  K   A  K  V  N   N  S  F   Q  N  T   I  P  F  N   I  F  S   Y  T  N   N  S  L  L   K  D  I   I  N  E   Y  F  N  N
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                2610       2620       2630       2640       2650       2660       2670       2680       2690       2700       2710       2720       2730
1 C-ST (Ref) tattaatgat tcaaaaattt tgagcctaca aaacagaaaa aatactttag tggatacatc aggatataat gcagaagtga gtgaagaagg cgatgttcag cttaatccaa tatttccatt tgactttaaa
2             I  N  D   S  K  I   L  S  L  Q   N  R  K   N  T  L   V  D  T  S   G  Y  N   A  E  V   S  E  E  G   D  V  Q   L  N  P   I  F  P  F   D  F  K
3 pGEM-Cst   tattaatgat tcaaaaattt tgagcctaca aaacagaaaa aatactttag tggatacatc aggatataat gcagaagtga gtgaagaagg cgatgttcag cttaatccaa tatttccatt tgactttaaa
4             I  N  D   S  K  I   L  S  L  Q   N  R  K   N  T  L   V  D  T  S   G  Y  N   A  E  V   S  E  E  G   D  V  Q   L  N  P   I  F  P  F   D  F  K
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                2740       2750       2760       2770       2780       2790       2800       2810       2820       2830       2840       2850       2860
1 C-ST (Ref) ttaggtagtt caggggagga tagaggtaaa gttatagtaa cccagaatga aaatattgta tataattcta tgtatgaaag ttttagcatt agtttttgga ttagaataaa taaatgggta agtaatttac
2             L  G  S   S  G  E  D   R  G  K   V  I  V   T  Q  N  E   N  I  V   Y  N  S   M  Y  E  S   F  S  I   S  F  W   I  R  I  N   K  W  V   S  N  L
3 pGEM-Cst   ttaggtagtt caggggagga tagaggtaaa gttatagtaa cccagaatga aaatattgta tataattcta tgtatgaaag ttttagcatt agtttttgga ttagaataaa taaatgggta agtaatttac
4             L  G  S   S  G  E  D   R  G  K   V  I  V   T  Q  N  E   N  I  V   Y  N  S   M  Y  E  S   F  S  I   S  F  W   I  R  I  N   K  W  V   S  N  L
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                2870       2880       2890       2900       2910       2920       2930       2940       2950       2960       2970       2980       2990
1 C-ST (Ref) ctggatatac tataattgat agtgttaaaa ataactcagg ttggagtata ggtattatta gtaatttttt agtatttact ttaaaacaaa atgaagatag tgaacaaagt ataaatttta gttatgatat
2            P  G  Y  T   I  I  D   S  V  K   N  N  S  G   W  S  I   G  I  I   S  N  F  L   V  F  T   L  K  Q   N  E  D  S   E  Q  S   I  N  F   S  Y  D  I
3 pGEM-Cst   ctggatatac Cataattgat agtgttaaaa ataactcagg ttggagtata ggtattatta gtaatttttt agtatttact ttaaaacaaa atgaagatag tgaacaaagt ataaatttta gttatgatat
4            P  G  Y  T   I  I  D   S  V  K   N  N  S  G   W  S  I   G  I  I   S  N  F  L   V  F  T   L  K  Q   N  E  D  S   E  Q  S   I  N  F   S  Y  D  I
5            ---------- *--------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                3000       3010       3020       3030       3040       3050       3060       3070       3080       3090       3100       3110       3120
1 C-ST (Ref) atcaaataat gctcctggat acaataaatg gtttttttgta actgttacta acaatatgat gggaaatatg aagatttata taaatggaaa attaatagat actataaaag ttaaagaact aactggaatt
2             S  N  N   A  P  G   Y  N  K  W   F  F  V   T  V  T   N  N  M  M   G  N  M   K  I  Y   I  N  G  K   L  I  D   T  I  K   V  K  E  L   T  G  I
3 pGEM-Cst   atcaaataat gctcctggat acaataaatg gtttttttgta actgttacta acaatatgat gggaaatatg aagatttata taaatggaaa attaatagat actataaaag ttaaagaact aactggaatt
4             S  N  N   A  P  G   Y  N  K  W   F  F  V   T  V  T   N  N  M  M   G  N  M   K  I  Y   I  N  G  K   L  I  D   T  I  K   V  K  E  L   T  G  I
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

```
                3130       3140       3150       3160       3170       3180       3190       3200       3210       3220       3230       3240       3250
1 C-ST (Ref) aattttagca aaactataac atttgaaata aataaaattc cagataccgg tttgattact tcagattctg ataacatcaa tatgtggata agagattttt atatatttgc taaagaatta gatggtaaag
2              N  F  S   K  T  I  T   F  E  I   N  K  I   P  D  T  G   L  I  T   S  D  S   D  N  I  N   M  W  I   R  D  F   Y  I  F  A   K  E  L   D  G  K
3 pGEM-Cst    aattttagca aaactataac atttgaaata aataaaattc cagataccgg tttgattact tcagattctg ataacatcaa tatgtggata agagattttt atatatttgc taaagaatta gatggtaaag
4              N  F  S   K  T  I  T   F  E  I   N  K  I   P  D  T  G   L  I  T   S  D  S   D  N  I  N   M  W  I   R  D  F   Y  I  F  A   K  E  L   D  G  K
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                3260       3270       3280       3290       3300       3310       3320       3330       3340       3350       3360       3370       3380
1 C-ST (Ref) atattaatat attatttaat agcttgcaat atactaatgt tgtaaaagat tattggggaa atgatttaag atataataaa gaatattata tggttaatat agattattta aatagatata tgtatgcgaa
2            D  I  N  I   L  F  N   S  L  Q   Y  T  N  V   V  K  D   Y  W  G   N  D  L  R   Y  N  K   E  Y  Y   M  V  N  I   D  Y  L   N  R  Y   M  Y  A  N
3 pGEM-Cst    atattaatat attatttaat agcttgcaat atactaatgt tgtaaaagat tattggggaa atgatttaag atataataaa gaatattata tggttaatat agattattta aatagatata tgtatgcgaa
4            D  I  N  I   L  F  N   S  L  Q   Y  T  N  V   V  K  D   Y  W  G   N  D  L  R   Y  N  K   E  Y  Y   M  V  N  I   D  Y  L   N  R  Y   M  Y  A  N
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                3390       3400       3410       3420       3430       3440       3450       3460       3470       3480       3490       3500       3510
1 C-ST (Ref) ctcacgacaa attgttttta atacacgtag aaataataat gacttcaatg aaggatataa aattataata aaaagaatca gaggaaatac aaatgatact agagtacgag gaggagatat tttatatttt
2              S  R  Q   I  V  F   N  T  R  R   N  N  N   D  F  N   E  G  Y  K   I  I  I   K  R  I   R  G  N  T   N  D  T   R  V  R   G  G  D  I   L  Y  F
3 pGEM-Cst    ctcacgacaa attgttttta atacacgtag aaataataat gacttcaatg aaggatataa aattataata aaaagaatca gaggaaatac aaatgatact agagtacgag gaggagatat tttatatttt
4              S  R  Q   I  V  F   N  T  R  R   N  N  N   D  F  N   E  G  Y  K   I  I  I   K  R  I   R  G  N  T   N  D  T   R  V  R   G  G  D  I   L  Y  F
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                3520       3530       3540       3550       3560       3570       3580       3590       3600       3610       3620       3630       3640
1 C-ST (Ref) gatatgacaa ttaataacaa agcatataat ttgtttatga agaatgaaac tatgtatgca gataatcata gtactgaaga tatatatgct ataggtttaa gagaacaaac aaaggatata aatgataata
2             D  M  T   I  N  N  K   A  Y  N   L  F  M   K  N  E  T   M  Y  A   D  N  H   S  T  E  D   I  Y  A   I  G  L   R  E  Q  T   K  D  I   N  D  N
3 pGEM-Cst    gatatgacaa ttaataacaa agcatataat ttgtttatga agaatgaaac tatgtatgca gataatcata gtactgaaga tatatatgct ataggtttaa gagaacaaac aaaggatata aatgataata
4             D  M  T   I  N  N  K   A  Y  N   L  F  M   K  N  E  T   M  Y  A   D  N  H   S  T  E  D   I  Y  A   I  G  L   R  E  Q  T   K  D  I   N  D  N
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                3650       3660       3670       3680       3690       3700       3710       3720       3730       3740       3750       3760       3770
1 C-ST (Ref) ttatatttca aatacaacca atgaataata cttattatta cgcatctcaa atatttaaat caaattttaa tggagaaaat atttctggaa tatgttcaat aggtacttat cgttttagac ttggaggtga
2            I  I  F  Q   I  Q  P   M  N  N   T  Y  Y  Y   A  S  Q   I  F  K   S  N  F  N   G  E  N   I  S  G   I  C  S  I   G  T  Y   R  F  R   L  G  G  D
3 pGEM-Cst    ttatatttca aatacaacca atgaataata cttattatta cgcatctcaa atatttaaat caaattttaa tggagaaaat atCtctggaa tatgttcaat aggtacttat cgttttagac ttggaggtga
4            I  I  F  Q   I  Q  P   M  N  N   T  Y  Y  Y   A  S  Q   I  F  K   S  N  F  N   G  E  N   I  S  G   I  C  S  I   G  T  Y   R  F  R   L  G  G  D
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- --*------- ---------- ---------- ---------- ----------

                3780       3790       3800       3810       3820       3830       3840       3850       3860       3870       3880       3890       3900
1 C-ST (Ref) ttggtataga cacaattatt tggtgcctac tgtgaagcaa ggaaattatg cttcattatt agaatcaaca tcaactcatt ggggtttttgt acctgtaagt gaataa
2              W  Y  R   H  N  Y   L  V  P  T   V  K  Q   G  N  Y   A  S  L  L   E  S  T   S  T  H   W  G  F  V   P  V  S   E  *
3 pGEM-Cst    ttggtataga cacaattatt tggtgcctac tgtgaagcaa ggaaattatg cttcattatt agaatcaaca tcaactcatt ggggtttttgt acctgtaagt gaataa
4              W  Y  R   H  N  Y   L  V  P  T   V  K  Q   G  N  Y   A  S  L  L   E  S  T   S  T  H   W  G  F  V   P  V  S   E  *
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------
```

Fig. 2.7 Comparison of complete nucleotide and deduced amino acid sequence of cloned BoNT/C_{ST} gene with the published sequence (Kimura *et al.*, 1990 (GI: 217780)). * Substituted bases are shown in capital letters.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

```
                      10         20         30         40         50         60         70         80         90        100        110        120        130
1 D-SA (Ref) atgacatggc cagtaaaaga ttttaattat agtgatcctg ttaatgacaa tgatatatta tatttaagaa taccacaaaa taagttaatt actacacctg taaaagcttt tatgattact caaaatattt
2             M  T  W   P  V  K  D   F  N  Y   S  D  P   V  N  D  N   D  I  L   Y  L  R   I  P  Q  N   K  L  I   T  T  P   V  K  A  F   M  I  T   Q  N  I
3 pGEM-Dsa   atgacatTgc cagtaaaaga ttttaattat agtgatcctg ttaatgacaa tgatatatta tatttaagaa taccacaaaa taagttaatt actacacctg taaaagcttt tatgattact caaaatattt
4             M  T  L   P  V  K  D   F  N  Y   S  D  P   V  N  D  N   D  I  L   Y  L  R   I  P  Q  N   K  L  I   T  T  P   V  K  A  F   M  I  T   Q  N  I
5            -------*-- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     140        150        160        170        180        190        200        210        220        230        240        250        260
1 D-SA (Ref) gggtaatacc agaaagattt tcatcagata ctaatccaag tttaagtaaa ccgcctagac ctacttcaaa gtatcaaagt tattatgatc ctagttattt atctactgat gagcaaaaag atacattttt
2            W  V  I  P   E  R  F   S  S  D   T  N  P  S   L  S  K   P  P  R   P  T  S  K   Y  Q  S   Y  Y  D   P  S  Y  L   S  T  D   E  Q  K   D  T  F  L
3 pGEM-Dsa   gggtaatacc agaaagattt tcatcagata ctaatccaag tttaagtaaa ccgcctagac ctacttcaaa gtatcaaagt tattatgatc ctagttattt atctactgat gagcaaaaag atacattttt
4            W  V  I  P   E  R  F   S  S  D   T  N  P  S   L  S  K   P  P  R   P  T  S  K   Y  Q  S   Y  Y  D   P  S  Y  L   S  T  D   E  Q  K   D  T  F  L
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     270        280        290        300        310        320        330        340        350        360        370        380        390
1 D-SA (Ref) aaaagggatt ataaaattat ttaaaagaat taatgaaaga gatataggaa aaaaattaat aaattattta gtagttggtt caccttttat gggagattca agtacgcctg aagatacatt tgattttaca
2              K  G  I   I  K  L   F  K  R  I   N  E  R   D  I  G   K  K  L  I   N  Y  L   V  V  G   S  P  F  M   G  D  S   S  T  P   E  D  T  F   D  F  T
3 pGEM-Dsa   aaaagggatt ataaaattat ttaaaagaat taatgaaaga gatataggaa aaaaattaat aaattattta gtagttggtt caccttttat gggagattca agtacgcctg aagatacatt tgattttaca
4              K  G  I   I  K  L   F  K  R  I   N  E  R   D  I  G   K  K  L  I   N  Y  L   V  V  G   S  P  F  M   G  D  S   S  T  P   E  D  T  F   D  F  T
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     400        410        420        430        440        450        460        470        480        490        500        510        520
1 D-SA (Ref) cgtcatacta ctaatattgc agttgaaaag tttgaaaatg gtagttggaa agtaacaaat attataacac caagtgtatt gatatttgga ccacttccta atatattaga ctatacagca tcccttacat
2             R  H  T   T  N  I  A   V  E  K   F  E  N   G  S  W  K   V  T  N   I  I  T   P  S  V  L   I  F  G   P  L  P   N  I  L  D   Y  T  A   S  L  T
3 pGEM-Dsa   cgtcatacta ctaatattgc agttgaaaag tttgaaaatg gtagttggaa agtaacaaat attataacac caagtgtatt gatatttgga ccacttccta atatattaga ctatacagca tcccttacat
4             R  H  T   T  N  I  A   V  E  K   F  E  N   G  S  W  K   V  T  N   I  I  T   P  S  V  L   I  F  G   P  L  P   N  I  L  D   Y  T  A   S  L  T
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     530        540        550        560        570        580        590        600        610        620        630        640        650
1 D-SA (Ref) tgcaaggaca acaatcaaat ccatcatttg aagggtttgg aacattatct atactaaaag tagcacctga atttttgtta acatttagtg atgtaacatc taatcaaagt tcagctgtat taggcaaatc
2            L  Q  G  Q   Q  S  N   P  S  F   E  G  F  G   T  L  S   I  L  K   V  A  P  E   F  L  L   T  F  S   D  V  T  S   N  Q  S   S  A  V   L  G  K  S
3 pGEM-Dsa   tgcaaggaca acaatcaaat ccatcatttg aagggtttgg aacattatct atactaaaag tagcacctga atttttgtta acatttagtg atgtaacatc taatcaaagt tcagctgtat taggcaaatc
4            L  Q  G  Q   Q  S  N   P  S  F   E  G  F  G   T  L  S   I  L  K   V  A  P  E   F  L  L   T  F  S   D  V  T  S   N  Q  S   S  A  V   L  G  K  S
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     660        670        680        690        700        710        720        730        740        750        760        770        780
1 D-SA (Ref) tatattttgt atggatccag taatagcttt aatgcatgag ttaacacatt ctttgcatca attgtatgga ataaatatac catctgataa aaggattcgt ccacaagtta gcgagggatt tttttctcaa
2              I  F  C   M  D  P   V  I  A  L   M  H  E   L  T  H   S  L  H  Q   L  Y  G   I  N  I   P  S  D  K   R  I  R   P  Q  V   S  E  G  F   F  S  Q
3 pGEM-Dsa   tatattttgt atggatccag taatagcttt aatgcatgag ttaacacatt ctttgcatca attgtatgga ataaatatac catctgataa aaggattcgt ccacaagtta gcgagggatt tttttctcaa
4              I  F  C   M  D  P   V  I  A  L   M  H  E   L  T  H   S  L  H  Q   L  Y  G   I  N  I   P  S  D  K   R  I  R   P  Q  V   S  E  G  F   F  S  Q
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     790        800        810        820        830        840        850        860        870        880        890        900        910
1 D-SA (Ref) gatggaccca acgtacaatt tgaggaatta tacacatttg gaggatcaga tgttgaaata atacctcaaa ttgaaagatt acaattaaga gaaaaagcat taggtcacta taaagatata gcgaaaagac
2             D  G  P   N  V  Q  F   E  E  L   Y  T  F   G  G  S  D   V  E  I   I  P  Q   I  E  R  L   Q  L  R   E  K  A   L  G  H  Y   K  D  I   A  K  R
3 pGEM-Dsa   gatggaccca acgtacaatt tgaggaatta tacacatttg gaggatcaga tgttgaaata atacctcaaa ttgaaagatt acaattaaga gaaaaagcat taggtcacta taaagatata gcgaaaagac
4             D  G  P   N  V  Q  F   E  E  L   Y  T  F   G  G  S  D   V  E  I   I  P  Q   I  E  R  L   Q  L  R   E  K  A   L  G  H  Y   K  D  I   A  K  R
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     920        930        940        950        960        970        980        990       1000       1010       1020       1030       1040
1 D-SA (Ref) ttaataatat taataaaact attccttcta gttggagtag taatatagat aaatataaaa aaatattttc tgaaaagtat aattttgata aagataatac aggaaatttt cttgtaaata ttgataaatt
2            L  N  N  I   N  K  T   I  P  S   S  W  S  S   N  I  D   K  Y  K   K  I  F  S   E  K  Y   N  F  D   K  D  N  T   G  N  F   L  V  N   I  D  K  F
3 pGEM-Dsa   ttaataatat taataaaact attccttcta gttggagtag taatatagat aaatataaaa aaatattttc tgaaaagtat aattttgata aagataatac aggaaatttt GCtgtaaata ttgataaatt
4            L  N  N  I   N  K  T   I  P  S   S  W  S  S   N  I  D   K  Y  K   K  I  F  S   E  K  Y   N  F  D   K  D  N  T   G  N  F   A  V  N   I  D  K  F
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- **-------- ----------
```

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

```
                   1050       1060       1070       1080       1090       1100       1110       1120       1130       1140       1150       1160       1170
1 D-SA (Ref) caatagctta tattcagact tgactaatgt tatgtcagaa gttatttatt cttcgcaata taatgttaaa aacaggactc attatttttc aaagcattat ctacctgtat ttgcaaatat attagatgat
2             N  S  L   Y  S  D   L  T  N  V   M  S  E   V  I  Y   S  S  Q  Y   N  V  K   N  R  T   H  Y  F  S   K  H  Y   L  P  V   F  A  N  I   L  D  D
3 pGEM-Dsa   caatagctta tattcagact tgactaatgt tatgtcagaa gttGtttatt cttcgcaata taatgttaaa aacaggactc attatttttc aaagcattat ctacctgtat ttgcaaatat attagatgat
4             N  S  L   Y  S  D   L  T  N  V   M  S  E   V  V  Y   S  S  Q  Y   N  V  K   N  R  T   H  Y  F  S   K  H  Y   L  P  V   F  A  N  I   L  D  D
5            ---------- ---------- ---------- ---------- ---*------ ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                   1180       1190       1200       1210       1220       1230       1240       1250       1260       1270       1280       1290       1300
1 D-SA (Ref) aatatttata ctataataaa cggttttaat ttaacaacta aaggttttaa tatagaaaat tcgggtcaga atatagaaag gaatcctgca ctacaaaaac ttagttcaga aagtgtagta gatttgttta
2             N  I  Y   T  I  I  N   G  F  N   L  T  T   K  G  F  N   I  E  N   S  G  Q   N  I  E  R   N  P  A   L  Q  K   L  S  S  E   S  V  V   D  L  F
3 pGEM-Dsa   aatatttata ctataataaa cggttttaat ttaacaacta aaggttttaa tatagaaaat tcgggCcaga atatagaaag gaatcctgca ctacaaaaac ttagttcaga aagtgtagta gatttgttta
4             N  I  Y   T  I  I  N   G  F  N   L  T  T   K  G  F  N   I  E  N   S  G  Q   N  I  E  R   N  P  A   L  Q  K   L  S  S  E   S  V  V   D  L  F
5            ---------- ---------- ---------- ---------- ---------- ---------- -----*---- ---------- ---------- ---------- ---------- ---------- ----------

                   1310       1320       1330       1340       1350       1360       1370       1380       1390       1400       1410       1420       1430
1 D-SA (Ref) caaaagtatg tttaagatta acaagaaata gtagagatga ttcaacatgt attcaagtta aaaataatac attaccttat gtagctgata aagatagcat ttcacaagaa atatttgaaa gtcaaattat
2            T  K  V  C   L  R  L   T  R  N   S  R  D  D   S  T  C   I  Q  V   K  N  N  T   L  P  Y   V  A  D   K  D  S  I   S  Q  E   I  F  E   S  Q  I  I
3 pGEM-Dsa   caaaagtatg tttaagatta acaagaaata gtagagatga ttcaacGtgt attcaagtta aaaataatac attaccttat gtagctgata aagatagcat ttcacaagaa atatttgaaa gtcaaattat
4            T  K  V  C   L  R  L   T  R  N   S  R  D  D   S  T  C   I  Q  V   K  N  N  T   L  P  Y   V  A  D   K  D  S  I   S  Q  E   I  F  E   S  Q  I  I
5            ---------- ---------- ---------- ---------- ------*--- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                   1440       1450       1460       1470       1480       1490       1500       1510       1520       1530       1540       1550       1560
1 D-SA (Ref) tacagatgag actaatgtag aaaattattc agataatttt tcattagatg aatctatttt agatgcaaaa gtccctacta atcctgaagc agtagatcca ctgttaccca atgttaatat ggaaccttta
2              T  D  E   T  N  V   E  N  Y  S   D  N  F   S  L  D   E  S  I  L   D  A  K   V  P  T   N  P  E  A   V  D  P   L  L  P   N  V  N  M   E  P  L
3 pGEM-Dsa   tacagatgag actaatgtag aaaattattc agataatttt tcattagatg aatctatttt agatgcaaaa gtccctacta atcctgaagc agtagatcca ctgttaccca atgttaatat ggaaccttta
4              T  D  E   T  N  V   E  N  Y  S   D  N  F   S  L  D   E  S  I  L   D  A  K   V  P  T   N  P  E  A   V  D  P   L  L  P   N  V  N  M   E  P  L
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                   1570       1580       1590       1600       1610       1620       1630       1640       1650       1660       1670       1680       1690
1 D-SA (Ref) aatgttccag gtgaagaaga agtattttat gatgatatta ctaaagatgt tgattattta aactcttatt attatttgga agcccaaaaa ttaagtaata atgttgaaaa tattactctt acaacttcag
2             N  V  P   G  E  E  E   V  F  Y   D  D  I   T  K  D  V   D  Y  L   N  S  Y   Y  Y  L  E   A  Q  K   L  S  N   N  V  E  N   I  T  L   T  T  S
3 pGEM-Dsa   aatgttccag gtgaagaaga agtattttat gatgatatta ctaaagatgt tgattattta aactcttatt attatttgga agcccaaaaa ttaagtaata atgttgaaaa tattactctt acaacttcag
4             N  V  P   G  E  E  E   V  F  Y   D  D  I   T  K  D  V   D  Y  L   N  S  Y   Y  Y  L  E   A  Q  K   L  S  N   N  V  E  N   I  T  L   T  T  S
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                   1700       1710       1720       1730       1740       1750       1760       1770       1780       1790       1800       1810       1820
1 D-SA (Ref) ttgaagaagc attaggttat agcaataaga tatacacatt tttacctagc ttagctgaaa aagtgaataa aggtgttcaa gcaggtttat tcttaaattg ggcgaatgaa gtagttgagg attttactac
2            V  E  E  A   L  G  Y   S  N  K   I  Y  T  F   L  P  S   L  A  E   K  V  N  K   G  V  Q   A  G  L   F  L  N  W   A  N  E   V  V  E   D  F  T  T
3 pGEM-Dsa   ttgaagaagc attaggttat agcaataaga tatacacatt tttacctagc ttagctgaaa aagtgaataa aggtgttcaa gcaggtttat tcttaaattg ggcgaatgaa gtagttgagg attttactac
4            V  E  E  A   L  G  Y   S  N  K   I  Y  T  F   L  P  S   L  A  E   K  V  N  K   G  V  Q   A  G  L   F  L  N  W   A  N  E   V  V  E   D  F  T  T
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                   1830       1840       1850       1860       1870       1880       1890       1900       1910       1920       1930       1940       1950
1 D-SA (Ref) aaatattatg aaaaaagata cattggataa aatatcagat gtatcagcca taattccata tataggacct gccttaaata taggaaattc agcattaagg ggaaacttta agcaagcatt tgcaacagct
2             N  I  M   K  K  D   T  L  D  K   I  S  D   V  S  A   I  I  P  Y   I  G  P   A  L  N   I  G  N  S   A  L  R   G  N  F   K  Q  A  F   A  T  A
3 pGEM-Dsa   aaatattatg aaaaaagata cattggataa aatatcagat gtatcagcca taattccata tataggacct gccttaaata taggaaattc agcattaagg ggaaacttta agcaagcatt tgcaacagct
4             N  I  M   K  K  D   T  L  D  K   I  S  D   V  S  A   I  I  P  Y   I  G  P   A  L  N   I  G  N  S   A  L  R   G  N  F   K  Q  A  F   A  T  A
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                   1960       1970       1980       1990       2000       2010       2020       2030       2040       2050       2060       2070       2080
1 D-SA (Ref) ggtgtagctt ttttgttaga aggatttcca gagtttacaa tacctgcact cggtgtattt accttttata gttctattca agaaagagag aaaattatta aaactataga aaattgttta gaacaaagag
2             G  V  A   F  L  L  E   G  F  P   E  F  T   I  P  A  L   G  V  F   T  F  Y   S  S  I  Q   E  R  E   K  I  I   K  T  I  E   N  C  L   E  Q  R
3 pGEM-Dsa   ggtgtagctt ttttgttaga aggatttcca gagtttacaa tacctgcact cggtgtattt accttttata gttctattca agaaagagag aaaattatta aaactataga aaattgttta gaacaaagag
4             G  V  A   F  L  L  E   G  F  P   E  F  T   I  P  A  L   G  V  F   T  F  Y   S  S  I  Q   E  R  E   K  I  I   K  T  I  E   N  C  L   E  Q  R
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

```
                 2090       2100       2110       2120       2130       2140       2150       2160       2170       2180       2190       2200       2210
1 D-SA (Ref) ttaagagatg gaaagattca tatcaatgga tggtatcaaa ttggttgtca agaattacta ctcaatttaa tcatataagt tatcaaatgt atgattcttt gagttatcag gcagatgcaa tcaaagctaa
2            V  K  R  W   K  D  S   Y  Q  W   M  V  S  N   W  L  S   R  I  T   T  Q  F  N   H  I  S   Y  Q  M   Y  D  S  L   S  Y  Q   A  D  A   I  K  A  K
3 pGEM-Dsa   ttaagagatg gaaagattca tatcaatgga tggtatcaaa ttggttgtca agaattacta ctcaatttaa tcatataagt tatcaaatgt atgattcttt gagttatcag gcagatgcaa tcaaagctaa
4            V  K  R  W   K  D  S   Y  Q  W   M  V  S  N   W  L  S   R  I  T   T  Q  F  N   H  I  S   Y  Q  M   Y  D  S  L   S  Y  Q   A  D  A   I  K  A  K
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                 2220       2230       2240       2250       2260       2270       2280       2290       2300       2310       2320       2330       2340
1 D-SA (Ref) aatagattta gaatataaaa aatactcagg aagtgataaa gaaaatataa aaagtcaagt tgaaaattta aaaaatagtt tagatgtaaa aatctcggaa gcaatgaata atataaataa atttatacga
2             I  D  L   E  Y  K   K  Y  S  G   S  D  K   E  N  I   K  S  Q  V   E  N  L   K  N  S   L  D  V  K   I  S  E   A  M  N   N  I  N  K   F  I  R
3 pGEM-Dsa   aatagattta gaatataaaa aatactcagg aagtgataaa gaaaatataa aaGgtcaagt tgaaaattta aaaaatagtt tagatgtaaa aatctcggaa gcaatgaata atataaataa atttatacga
4             I  D  L   E  Y  K   K  Y  S  G   S  D  K   E  N  I   K  G  Q  V   E  N  L   K  N  S   L  D  V  K   I  S  E   A  M  N   N  I  N  K   F  I  R
5            ---------- ---------- ---------- ---------- ---------- --*------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                 2350       2360       2370       2380       2390       2400       2410       2420       2430       2440       2450       2460       2470
1 D-SA (Ref) gaatgttctg taacatactt atttaaaaat atgctcccta aagtaattga tgaattaaat aagtttgatt taaaaactaa aacagaatta attaatctta tagatagtca taatattatt ctagttggtg
2             E  C  S   V  T  Y  L   F  K  N   M  L  P   K  V  I  D   E  L  N   K  F  D   L  K  T  K   T  E  L   I  N  L   I  D  S  H   N  I  I   L  V  G
3 pGEM-Dsa   gaatgttctg taacatactt atttaaaaat atgctcccta aagtaattga tgaattaaat aagtttgatt taaaaactaa aacagaatta attaatctta tagatagtca taatattatt ctagttggtg
4             E  C  S   V  T  Y  L   F  K  N   M  L  P   K  V  I  D   E  L  N   K  F  D   L  K  T  K   T  E  L   I  N  L   I  D  S  H   N  I  I   L  V  G
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                 2480       2490       2500       2510       2520       2530       2540       2550       2560       2570       2580       2590       2600
1 D-SA (Ref) aagtagatag attaaaagca aaagtaaatg agagttttga aaatacaata ccctttaata ttttttcata tactaataat tctttattaa aagatatgat taatgaatat ttcaatagta ttaatgattc
2            E  V  D  R   L  K  A   K  V  N   E  S  F  E   N  T  I   P  F  N   I  F  S  Y   T  N  N   S  L  L   K  D  M  I   N  E  Y   F  N  S   I  N  D  S
3 pGEM-Dsa   aagtagatag attaaaagca aaagtaaatg agagttttga aaatacaata ccctttaata ttttttcata tactaataat tctttattaa aagatatgat taatgaatat ttcaatagta ttaatgattc
4            E  V  D  R   L  K  A   K  V  N   E  S  F  E   N  T  I   P  F  N   I  F  S  Y   T  N  N   S  L  L   K  D  M  I   N  E  Y   F  N  S   I  N  D  S
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                 2610       2620       2630       2640       2650       2660       2670       2680       2690       2700       2710       2720       2730
1 D-SA (Ref) aaaaattttg agcttacaaa ataaaaaaaa tactttgatg gatacatcag gatataacgc agaagtgaga gtagaaggca atgttcagct taatccaata tttccatttg actttaaatt aggtagttca
2             K  I  L   S  L  Q   N  K  K  N   T  L  M   D  T  S   G  Y  N  A   E  V  R   V  E  G   N  V  Q  L   N  P  I   F  P  F   D  F  K  L   G  S  S
3 pGEM-Dsa   aaaaattttg agcttacaaa ataaaaaaaa tactttgatg gatacatcag gatataacgc agaagtgaga gtagaaggca atgttcagct taatccaata tttccatttg actttaaatt aggtagttca
4             K  I  L   S  L  Q   N  K  K  N   T  L  M   D  T  S   G  Y  N  A   E  V  R   V  E  G   N  V  Q  L   N  P  I   F  P  F   D  F  K  L   G  S  S
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                 2740       2750       2760       2770       2780       2790       2800       2810       2820       2830       2840       2850       2860
1 D-SA (Ref) ggggatgata gaggtaaaat tatagtaacc cagaatgaaa atattgtata taatgctatg tatgaaagtt ttagtattag tttttggatt aggataaata aatgggtaag taatttacct ggatatacta
2             G  D  D   R  G  K  I   I  V  T   Q  N  E   N  I  V  Y   N  A  M   Y  E  S   F  S  I  S   F  W  I   R  I  N   K  W  V  S   N  L  P   G  Y  T
3 pGEM-Dsa   ggggatgata gaggtaaaGt tatagtaacc cagaatgaaa atattgtata taatgctatg tatgaaagtt ttagtattag tttttggatt aggataaata aatgggtaag taatttacct ggatatacta
4             G  D  D   R  G  K  V   I  V  T   Q  N  E   N  I  V  Y   N  A  M   Y  E  S   F  S  I  S   F  W  I   R  I  N   K  W  V  S   N  L  P   G  Y  T
5            ---------- --------*- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                 2870       2880       2890       2900       2910       2920       2930       2940       2950       2960       2970       2980       2990
1 D-SA (Ref) taattgatag tgttaaaaat aactcaggtt ggagtatagg tattattagt aattttttag tgtttacttt aaaacaaaat gaaaatagtg aacaagatat aaactttagt tatgatatat caaagaatgc
2            I  I  D  S   V  K  N   N  S  G   W  S  I  G   I  I  S   N  F  L   V  F  T  L   K  Q  N   E  N  S   E  Q  D  I   N  F  S   Y  D  I   S  K  N  A
3 pGEM-Dsa   taattgatag tgtCaaaaat aactcaggtt ggagtatagg tattattagt aattttttag tgtttacttt aaaacaaaat gaaaatagtg aacaagatat aaactttagt tatgatatat caaagaatgc
4            I  I  D  S   V  K  N   N  S  G   W  S  I  G   I  I  S   N  F  L   V  F  T  L   K  Q  N   E  N  S   E  Q  D  I   N  F  S   Y  D  I   S  K  N  A
5            ---------- ---*------ ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                 3000       3010       3020       3030       3040       3050       3060       3070       3080       3090       3100       3110       3120
1 D-SA (Ref) tgcgggatat aataaatggt tttttgtaac tattactacc aatatgatgg gaaatatgat gatttatata aatggaaaat taatagatac tataaaagtt aaagagttaa ctggaattaa ttttagcaaa
2             A  G  Y   N  K  W   F  F  V  T   I  T  T   N  M  M   G  N  M  M   I  Y  I   N  G  K   L  I  D  T   I  K  V   K  E  L   T  G  I  N   F  S  K
3 pGEM-Dsa   tgcgggatat aataaatggt tttttgtaac tattactacc aatatgatgg gaaatatgat gatttatata aatggaaaat taatagatac tataaaagtt aaagagttaa ctggaattaa ttttagcaaa
4             A  G  Y   N  K  W   F  F  V  T   I  T  T   N  M  M   G  N  M  M   I  Y  I   N  G  K   L  I  D  T   I  K  V   K  E  L   T  G  I  N   F  S  K
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

```
                     3130       3140       3150       3160       3170       3180       3190       3200       3210       3220       3230       3240       3250
1 D-SA (Ref) actataacat ttcaaatgaa taaaattcca aatactggct taattacctc agattctgat aacatcaata tgtggataag ggatttttat atctttgcta aagaattaga tgataaagat attaatatat
2             T  I  T   F  Q  M  N   K  I  P   N  T  G   L  I  T  S   D  S  D   N  I  N   M  W  I  R   D  F  Y   I  F  A   K  E  L  D   D  K  D   I  N  I
3 pGEM-Dsa   actGtaacat ttcaaatgaa taaaattcca aatactggct taattacctc agattctgat aacatcaata tgtggataag ggatttttat atctttgcta aagaattaga tgataaagat attaatatat
4             T  V  T   F  Q  M  N   K  I  P   N  T  G   L  I  T  S   D  S  D   N  I  N   M  W  I  R   D  F  Y   I  F  A   K  E  L  D   D  K  D   I  N  I
5            ---*------ ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     3260       3270       3280       3290       3300       3310       3320       3330       3340       3350       3360       3370       3380
1 D-SA (Ref) tatttaatag cttgcaatat actaatgttg taaaagatta ttggggaaat gatttaagat atgataaaga atattacatg attaacgtaa attatatgaa tagatatatg tctaaaaaag gcaatggaat
2            L  F  N  S   L  Q  Y   T  N  V   V  K  D  Y   W  G  N   D  L  R   Y  D  K  E   Y  Y  M   I  N  V   N  Y  M  N   R  Y  M   S  K  K   G  N  G  I
3 pGEM-Dsa   tatttaatag cttgcaatat actaatgttg taaaagatta ttggggaaat gatttaagat atgataaaga atattacatg attaacgtaa attatatgaa tagatatatg tctaaaaaag gcaatggaat
4            L  F  N  S   L  Q  Y   T  N  V   V  K  D  Y   W  G  N   D  L  R   Y  D  K  E   Y  Y  M   I  N  V   N  Y  M  N   R  Y  M   S  K  K   G  N  G  I
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     3390       3400       3410       3420       3430       3440       3450       3460       3470       3480       3490       3500       3510
1 D-SA (Ref) tgtttttaat acacgtaaaa ataataatga cttcaatgaa ggatataaaa ttataataaa aagaattaga ggaaatacaa atgatactag agtacgagga gaaaatgtat tatattttaa tactacaatt
2             V  F  N   T  R  K   N  N  N  D   F  N  E   G  Y  K   I  I  I  K   R  I  R   G  N  T   N  D  T  R   V  R  G   E  N  V   L  Y  F  N   T  T  I
3 pGEM-Dsa   tgtttttaat acacgtaaaa ataataatga cttcaatgaa ggatataaaa ttataataaa aagaattaga ggaaatacaa atgatactag agtacgagga gaaaatgtat tatattttaa tactacaatt
4             V  F  N   T  R  K   N  N  N  D   F  N  E   G  Y  K   I  I  I  K   R  I  R   G  N  T   N  D  T  R   V  R  G   E  N  V   L  Y  F  N   T  T  I
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     3520       3530       3540       3550       3560       3570       3580       3590       3600       3610       3620       3630       3640
1 D-SA (Ref) gataacaaac aatatagttt aggtatgtat aaaccttcta gaaatctagg gactgattta gttccactag gtgcattgga tcaaccaatg gatgagatac gtaaatatgg ttcgtttata atacaaccat
2             D  N  K   Q  Y  S  L   G  M  Y   K  P  S   R  N  L  G   T  D  L   V  P  L   G  A  L  D   Q  P  M   D  E  I   R  K  Y  G   S  F  I   I  Q  P
3 pGEM-Dsa   gataacaaac aatatagttt aggtatgtat aaaccttcta gaaatctagg gactgattta gttccactag gtgcattgga tcaaccaatg gatgagatac gtaaatatgg ttcgtttata atacaaccat
4             D  N  K   Q  Y  S  L   G  M  Y   K  P  S   R  N  L  G   T  D  L   V  P  L   G  A  L  D   Q  P  M   D  E  I   R  K  Y  G   S  F  I   I  Q  P
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     3650       3660       3670       3680       3690       3700       3710       3720       3730       3740       3750       3760       3770
1 D-SA (Ref) gcaatacttt tgattactat gcatcacaat tatttttgtc aagtaatgca acaacaaata ggcttggaat actatcaatt ggtagttata gtttcaaact tggagatgac tattggttta atcacgaata
2            C  N  T  F   D  Y  Y   A  S  Q   L  F  L  S   S  N  A   T  T  N   R  L  G  I   L  S  I   G  S  Y   S  F  K  L   G  D  D   Y  W  F   N  H  E  Y
3 pGEM-Dsa   gcaatacttt tgattactat gcatcacaat tatttttgtc aagtaatgca acaacaaata ggcttggaat actatcaatt ggtagttata gtttcaaact tggagatgac tattggttta atcacgaata
4            C  N  T  F   D  Y  Y   A  S  Q   L  F  L  S   S  N  A   T  T  N   R  L  G  I   L  S  I   G  S  Y   S  F  K  L   G  D  D   Y  W  F   N  H  E  Y
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

                     3780       3790       3800       3810       3820       3830       3840       3850       3860       3870       3880       3890       3900
1 D-SA (Ref) tttaattcct gttataaaaa tagagcatta tgcttcatta ttagaatcaa catcaactca ttgggttttt gtacctgcaa gtgaataa
2             L  I  P   V  I  K   I  E  H  Y   A  S  L   L  E  S   T  S  T  H   W  V  F   V  P  A   S  E  *
3 pGEM-Dsa   tttaattcct gttataaaaa tagagcatta tgcttcatta ttagaatcaa catcaactca ttgggttttt gtacctgTaa gtgaataa
4             L  I  P   V  I  K   I  E  H  Y   A  S  L   L  E  S   T  S  T  H   W  V  F   V  P  V   S  E  *
5            ---------- ---------- ---------- ---------- ---------- ---------- ---------- -------*-- --------
```

Fig. 2.8 Comparison of complete nucleotide and derived amino acid sequence of cloned BoNT/D$_{SA}$ gene with the published sequence (Moriishi *et al.*, 1996(a) (GI: 1374775)). * Substituted bases are shown in capital letters.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

2.3.2 Construction of mutant clones

Inverse PCR was used for the site-directed mutagenesis of the zinc-binding region contained in the botulinum neurotoxin encoding genes. Substitution of six nucleotides resulted in the mutation of three essential amino acids in the unique zinc-binding motif from His-Glu-aa-aa-His to Gly-Thr-aa-aa-Asn. Oligonucleotide primers were designed in an inverted tail-to-tail direction with nucleotides specifying *KpnI* restriction sites at their 5' ends. Using this PCR approach the vector DNA as well as the cloned gene was amplified. The choice of the *KpnI* site was motivated by the observation that neither pGEM®-T Easy nor the two BoNT genes contained such a restriction enzyme site. Following agarose gel electrophoresis, amplicons of ca. 6.8 kb were observed (Fig. 2.9). The size of the amplicons corresponds to the expected size of the plasmid DNA (3 kb) together with the BoNT genes (ca. 3.8 kb).

The purified amplicons were digested with *KpnI* to create compatible cohesive ends. The fragments were subsequently gel-purified and then self-ligated before being transformed into competent *E. coli* JM109 cells. The derived recombinant plasmids were characterized by restriction enzyme analysis. *KpnI* digestion linearized the recombinant plasmids and generated a 6.8 kb restriction fragment (Fig. 2.11), indicating that the respective recombinant plasmids contained the *KpnI* site and that the desired substitutions were indeed generated. Restriction of the recombinant plasmids by *SalI* (Fig. 2.11) resulted in two fragments being observed, namely a fragment corresponding to the 3.8 kb BoNT genes and a fragment corresponding to the size of the 3.0 kb pGEM®-T Easy vector.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

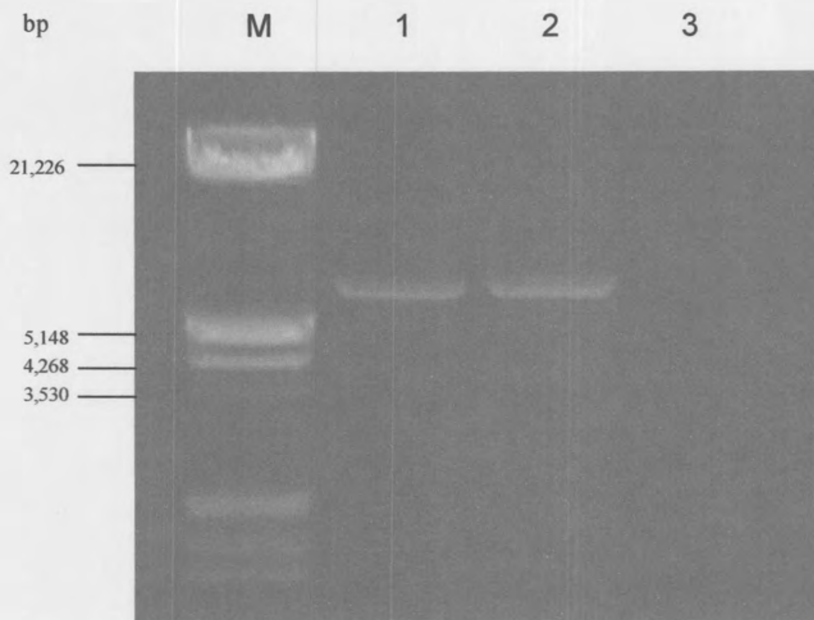

Fig. 2.9 Agarose gel electrophoresis indicating amplification of the pGEM®-T Easy vector together with the BoNT/C_{ST} and BoNT/D_{SA} genes. Lane: M, λ DNA/*EcoRI* + *HindIII* Marker; 1, pGEM-BoNT/C_{ST} amplicon; 2, pGEM-BoNT/D_{SA} amplicon; 3, Negative control.

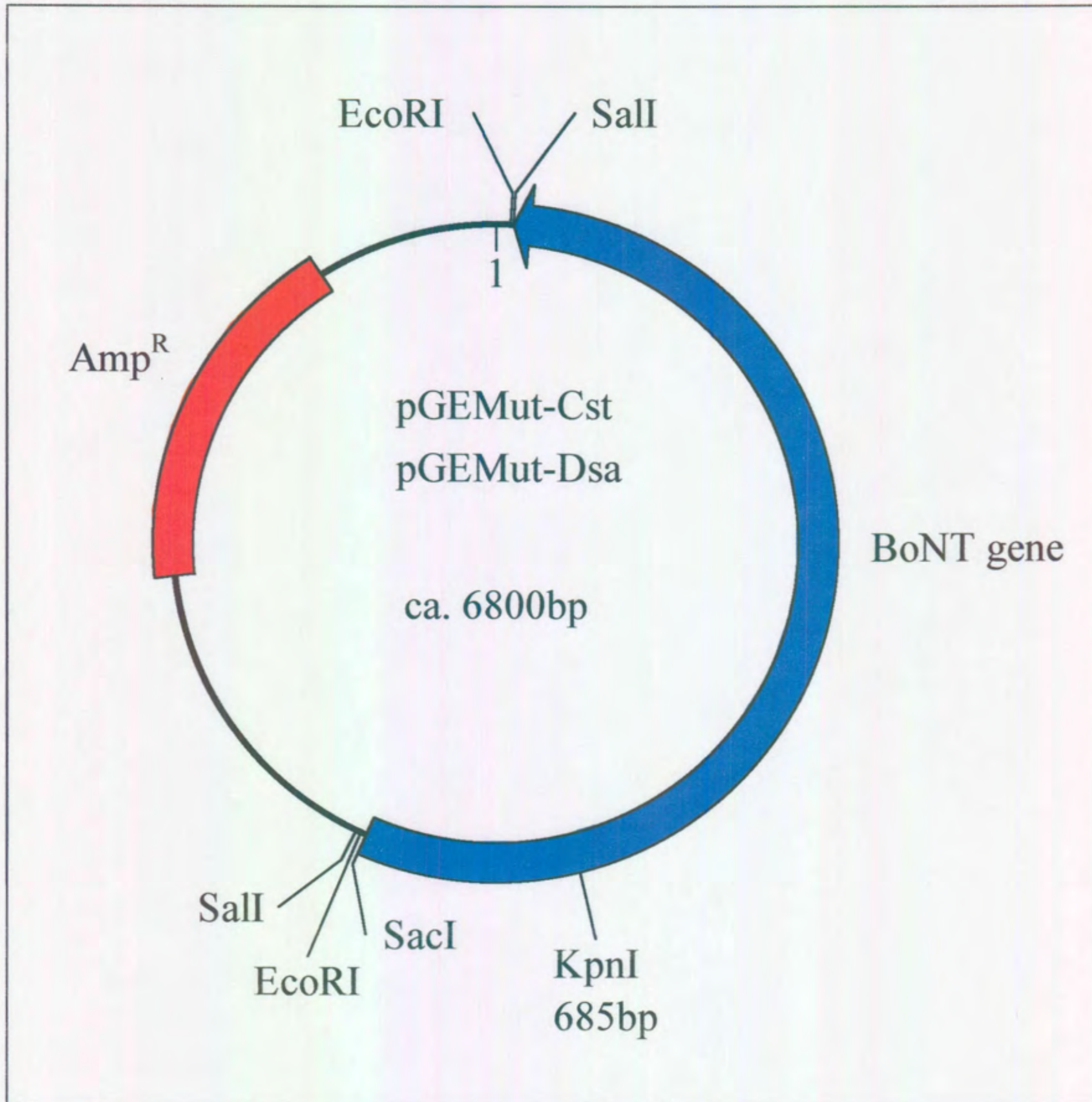

Fig. 2.10 Partial restriction enzyme map of recombinant pGEM®-T Easy vector containing either the mutated BoNT/C_{ST} or BoNT/D_{SA} gene.

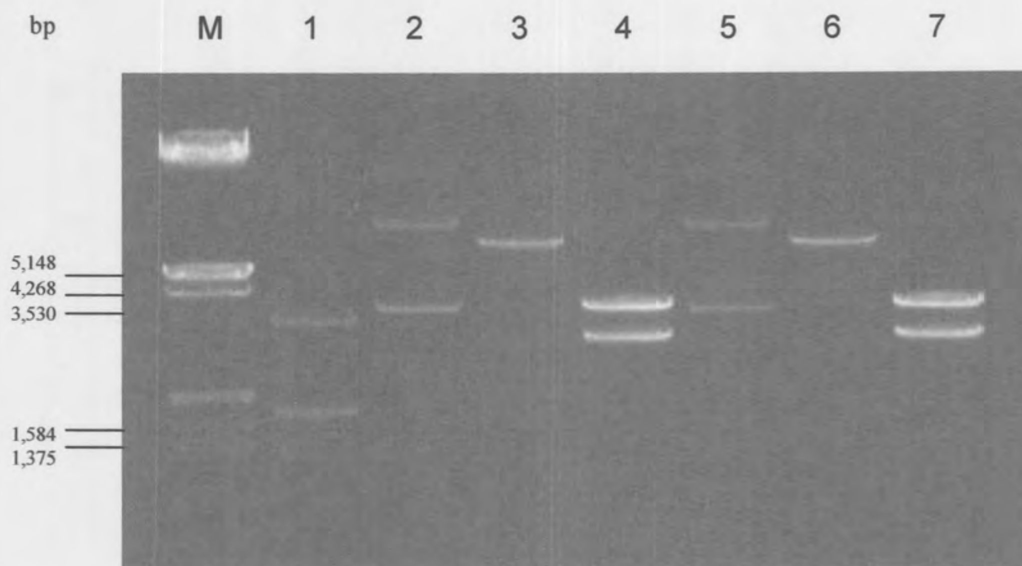

Fig. 2.11 Agarose gel electrophoretic analysis of the recombinant pGEM®-T Easy vector containing either the C_{ST} or D_{SA} genes, obtained by self-ligation of the mutagenized PCR amplicons prior to transformation of competent cells. Lanes: M, λ DNA/*EcoRI* + *HindIII* Marker; 1, Uncut non-recombinant pGEM®-T Easy vector; 2, Uncut recombinant type C_{ST} plasmid; 3, *KpnI*-digested recombinant type C_{ST} plasmid; 4, *SalI*-digested recombinant type C_{ST} plasmid; 5, Uncut recombinant type D_{SA} plasmid; 6, *KpnI*-digested recombinant type D_{SA} plasmid; 7, *SalI*-digested recombinant type D_{SA} plasmid.

In order to verify the integrity of the desired substitutions, approximately 600 nucleotides were analyzed with automated sequencing using BNTseqC4 and BNTseqD4 primers (Table 2.2) for the respective genes. Analysis of the obtained sequence indicated that the sequence contained in the zinc-binding motif (Table 2.5) was substituted by nucleotides specifying three different amino acid residues (Gly-Thr-aa-aa-Asn). No other differences were observed in the region sequenced. Plasmids containing the desired mutations were therefore designated pGEMut-C_{ST} and pGEMut-D_{SA}, respectively.

Table 2.5 Nucleotide and deduced amino acid sequences of zinc-binding motif of wild-type and mutant type C_{ST} and D_{SA} botulinum neurotoxin genes.

BoNT gene	**Nucleotide sequence**	**Amino acid sequence**
Wild-type C_{ST}	5'-CATGAACTTAATCAT-3'	His-Glu-Leu-Asn-His
Mutant C_{ST}	5'-**ggTacc**CTTAAT**a**AT-3'*	***Gly-Thr***-Leu-Asn-***Asn***§
Wild-type D_{SA}	5'-CATGAGTTAACACAT-3'	His-Glu-Leu-Thr-His
Mutant D_{SA}	5'-**ggTacc**TTAACA**a**AT-3'*	***Gly-Thr***-Leu-Thr-***Asn***§

* Substituted bases are shown in bold lower case letters.

§ Amino acid substitutions are shown in bold *italics*.

2.4 DISCUSSION

In order to express the mutant *Clostridium botulinum* neurotoxin genes in an eukaryotic baculovirus expression system, it was necessary to first clone, sequence and mutate the respective BoNT genes.

In 1968 Inòue and Iida illustrated that strains of *C. botulinum* carry bacteriophages (Inoue and Iida, 1968). The type C and D phages share more similarities than do any of the other *C. botulinum* phages and strains of these types are generally associated with

large, octahedral headed phages with long contractile tails with sheaths. A direct relationship between the type C and D bacteriophages and toxicogenicity has been established (Eklund *et al.*, 1971). Furthermore, it has been reported that toxin production by these cells declines rapidly and eventually ceases, following laboratory passaging of the bacterial cultures. During experimentation, bacteriophage loss was demonstrated with TEM in the type C_{ST} culture and this subsequently resulted in the lack of the appropriate target DNA for PCR amplification. Therefore, new type C_{ST} cultures had to be obtained. Thereafter, bacteriophages fitting the description of the type C and D phages were demonstrated with TEM.

The botulinum neurotoxin genes were amplified by using polymerase chain reaction (PCR) amplification. This is the first direct amplification of the full-length BoNT/C_{ST} and BoNT/D_{SA} genes. Previous reports have described the reconstruction of these genes with PCR fragments generated from culture supernatants and phage DNA libraries (Kimura *et al.*, 1990; Moriishi *et al.*, 1996a,b). One of the problems encountered during the amplification of the full-length genes involved the use of traditional Taq DNA polymerase enzymes. Taq DNA polymerase (Promega) and AmpliTaq Gold (Perkin-Elmer Biosystems) were unable to amplify the full-length botulinum neurotoxin genes. This can possibly be due to the length (ca. 3.8 kb) and high A-T (ca. 72%) content of the BoNT genes. Subsequently, a long-range enzyme specifically adapted for long and accurate amplification of DNA was used. The amplicons were cloned directly into the pGEM®-T Easy vector, since most PCR products amplified with TaKaRa Ex Taq™ have one A added at the 3' terminus. A ddT-tailed vector, such as pGEM®-T Easy, can ligate directly to an amplicon that contains a 3' overhanging A by forming a phoshodiester bond between the 5'-phosphate of the vector and the 3'-hydroxyl group from the amplicons' overhanging A (Holton and Graham, 1991).

The integrity of the full-length botulinum neurotoxin genes was subsequently verified by automated sequencing. The nucleotide sequences were compared with the published sequences for the respective genes and homologies of 99.8% (BoNT/C_{ST}) and 99.7% (BoNT/D_{SA}) were observed. Deduced amino acid sequences showed an open reading frame without fortuitous termination signals. Five amino acid dissimilarities were observed between the cloned BoNT/C_{ST} gene and published sequences (Kimura *et al.*,

1990). Similarly, 7 amino acid dissimilarities were observed between the cloned BoNT/D_{SA} and the published sequences (Moriishi *et al.,* 1996a).

Botulinum neurotoxins are zinc metallo-proteases with an unique zinc-binding motif (His-Glu-aa-aa-His) in the central region of the light chain sub-unit polypeptide. The zinc-protease activity of these neurotoxins causes the blockage of neurotransmitter release resulting in flaccid paralysis (Fu *et al.*, 1998). Kiyatkin *et al.* (1997) illustrated that the substitution of the three amino acids essential for zinc binding resulted in the complete detoxification of the modified recombinant type C toxin. The non-toxic modified type C toxin retained the ability to translocate from the gut to the general circulation and to evoke protective antibodies. This molecule may be an ideal vaccine candidate for botulism.

In this study, modified BoNT/C_{ST} and BoNT/D_{SA} genes were generated with site-directed mutagenesis of the three essential amino acids (His-Glu-aa-aa-His) in the zinc-binding region. The cloned full-length wild-type neurotoxin gene was subsequently used as template DNA in a PCR-based mutagenesis procedure. Mutations was generated by inverse PCR, in which a pair of inverted tail-to-tail primers were designed from the site of the desired mutation and used to amplify the entire double-stranded plasmid containing the desired gene. A unique *KpnI* restriction site was incorporated into the 5' ends of both the primers to introduce base pair substitutions and thus resulting in mutation of two of the three amino acids of importance. The *KpnI* site has the added advantages of creating compatible cohesive ends in the PCR product for increased circularization efficiency and it supports the fast and cost-effective screening of mutants. Restriction enzyme analysis and sequencing data of the mutated region indicated that the zinc binding region was successfully changed from His-Glu-aa-aa-His to Gly-Thr-aa-aa-Asn.

In this part of the project, the full-length wild-type BoNT/C_{ST} and BoNT/D_{SA} genes were successfully amplified, cloned and mutated. Sequencing analysis verified the integrity of the full-length wild-type genes and newly introduced mutations. These wild-type and mutant genes are of importance in further expression studies towards the development of a recombinant vaccine candidate for botulism.

CHAPTER 3

EXPRESSION OF WILD-TYPE AND MUTANT NEUROTOXIN GENES OF *Clostridium botulinum* TYPE C AND D IN THE BAC-TO-BAC™ BACULOVIRUS EXPRESSION SYSTEM

3.1 INTRODUCTION

The botulinum neurotoxins are one of the most toxic substances known to man. They are secreted as a single polypeptide chain with a molecular mass of approximately 150 kDa. The polypeptide is proteolytically cleaved to give rise to a light chain (L, 50 kDa) and a heavy chain (H, 100 kDa) that are linked by a disulfide bond. The L domain is responsible for enzymatic modification of protein components of the neuroexocytosis apparatus. The H chain can be divided into two domains (H_C and H_N) corresponding to the carboxy- and amino-terminal halves of the H chain, respectively. The H_C domain seems to be mainly responsible for neurospecific binding to the presynaptic membrane, while the H_N domain appears to govern internalization and membrane translocation.

Despite recent advances in gene expression systems, *E. coli* has remained a popular host organism for the prokaryotic expression of foreign proteins. Important advantages of using *E. coli* in expression studies include, among others, the rapid generation of biomass due to high rates of cell growth and the availability of low-cost culture conditions (Balbas and Bolivar, 1990). However, *E. coli* has a limited capacity to secrete proteins and it is unable to exert certain post-translational modifications on expressed foreign proteins e.g. disulfide bond formation, glycosylation, and acetylation. This usually leads to improper folding and consequently insolubility of the expressed protein.

Various reports have described expression of the botulinum neurotoxin (BoNT) and tetanus neurotoxin (TeNT) genes in *E. coli.* The tetanus neurotoxin closely resembles the botulinum neurotoxins in structure and biological activity. Makoff *et al.* (1989) have reported difficulties in obtaining high expression levels for the native gene fragment encoding the H_C domain of TeNT in *E. coli.* Their study indicated that this might be due to the presence of rare codons in the native gene of *C. tetani.* Expression of the native

gene fragment encoding BoNT/A(H_C) in *E. coli* resulted in similar low levels of expression (LePenotiere *et al.,* 1995). In order to overcome these low expression levels in *E. coli*, a synthetic BoNT/A(Hc) gene fragment was constructed by taking into account the single codon preferences of *E. coli*. Some of the codons were also altered to less preferred triplets in order to reduce the adenosine (A) and thymine (T) base content of the gene fragment, and to incorporate restriction enzyme recognition sites, which enabled construction of the synthetic gene. Expression of this gene construct yielded a low level of immunoreactive material that was largely insoluble (Clayton *et al.,* 1995). Kiyatkin *et al.* (1997) reported the successful expression (1 to 2 mg/l) of a full-length mutant BoNT/C gene in *E. coli*. The mutant BoNT/C gene was generated by site-directed mutagenesis of three amino acids within the zinc-binding motif of the enzymatic domain. The protein was synthesized in a soluble form and protective immunity was evoked by the oral and subcutaneous administration of this non-toxic mutant BoNT/C.

The use of *Saccharomyces cerevisiae* and *Pichia pastoris* as a means to express the tetanus and botulinum neurotoxins has also been reported. These eukaryotic gene expression systems are capable of secreting and performing post-translational modifications on expressed proteins. However, expression of the native gene fragment encoding the H_C domain of TeNT in *S. cerevisiae* has proved to be even more troublesome than in *E. coli*. Romanos *et al.* (1991) reported that the high adenosine and thymine base composition found in the native gene of clostridial DNA gives rise to fortuitous transcriptional termination signals in yeast. This obstacle was overcome by preparing a synthetic gene in which codon usage was altered to reduce the A+T content of the gene fragment (Romanos *et al.*, 1991). Intracellular expression of the synthetic gene fragment encoding TeNT(H_C) resulted in a soluble immunoreactive protein (50 kDa) with yields of 60 to 90 mg/l. Extracellular expression yielded hyper-glycosylated (75 to 200 kDa) and core-glycosylated (65 kDa) forms of the protein with a much lower yield of 5 to 10 mg/l. The secreted glycosylated proteins failed to elicit protective immunity in mice, but following a deglycosylation step their capacity to elicit an immune response was restored. Romanos *et al.* (1991) suggested that the added carbohydrates must sterically hinder the generation of neutralizing antibodies to adjacent epitopes. Smith (1998) has reported the successful expression of synthetic gene fragments encoding BoNT/A(H_C) and BoNT/B(H_C) in *Pichia pastoris* as both intra- and extracellular proteins. Intracellular expression of both genes yielded immunoreactive products with

the expected molecular weight as judged by SDS-PAGE and Western blot analysis. In contrast, the secreted proteins were found to have a higher molecular weight due to glycosylation of the proteins. However, only the BoNT/A(H_C) protein proved to be immunogenic, while the glycosylated BoNT/B(Hc) failed to elicit protective immunity in mice. Similar to TeNT(H_C), the capacity of BoNT/B(H_C) to protect mice from challenge with native toxin was restored only after deglycosylation of the protein.

From the preceding sections, it would thus appear that expression of genes derived from BoNT in *E. coli* occurs at variable levels and is generally inefficient. In addition *E. coli* has a disadvantage as a host organism in that it contains toxic cell wall pyrogens that cause fever in humans and animals. Furthermore, expression of synthetic gene fragments encoding the H_C domain in different yeast expression systems, has also been met with mixed results. Intracellular expression yielded high levels of TeNT(H_C), BoNT/A(H_C), and BoNT/B(H_C) which was shown to be immunogenic. In contrast, extracellular expression yielded glycosylated forms of the proteins at much lower levels and only BoNT/A(H_C) proved to be immunogenic. In order to obtain immunogenic extracellularly expressed proteins, either a deglycosylation step can be incorporated into the manufacturing process or the potential glycosylation site(s) in the protein can be altered by site-directed mutagenesis. The most cost-effective alternative, however, would be to express the proteins intracellularly in order to avoid the secretory pathway altogether (Smith, 1998).

The baculovirus expression system is an alternative eukaryotic expression system that may prove to be useful for expression of the BoNT genes. Successful expression of genes from various sources by baculovirus recombinants have been reported (O'Reilly *et al.,* 1992 and references therein). Not only does it provide an eukaryotic environment that is generally conducive to the proper folding, disulfide bond formation, oligomerization, and/or post-translational modification required for biological activity of some proteins (O'Reilly *et al.*, 1992), but biologically active proteins are synthesized more reliably than in bacterial or yeast expression systems, and the yields generally exceed those from mammalian expression systems (Liebman *et al.*, 1999).

The aim of this study was thus to express the full-length wild-type and mutant BoNT/C_{ST} and BoNT/D_{SA} genes in *S. frugiperda* cells using the BAC-TO-BAC™ baculovirus system. This will allow for further studies to evaluate the effectiveness of the introduced mutation for the inactivation of the neurotoxins as well as to evaluate the potential of the expressed proteins as recombinant vaccine candidates.

3.2 MATERIALS AND METHODS

3.2.1 Materials

In the previous chapter the construction of plasmids containing the wild-type and mutant type C and D neurotoxin genes was described (pGEM-C_{ST}, pGEM-D_{SA}, pGEMut-C_{ST}, and pGEMut-D_{SA}). An *E. coli* strain DH10Bac™ containing a vendor-supplied baculovirus shuttle vector (bacmid: Life technologies) was used for transformations with the recombinant donor pFASTBAC™DUAL plasmids (Life Technologies Ltd., Paisley, UK).

Rabbit anti-sera against the L chain of BoNT/C_{ST} (anti-C_{ST}(L)) and full-length BoNT/D_{1873} (anti-D_{1873}) were obtained from K. Oguma (Department of Bacteriology, Okayama University Medical School, Okayama, Japan)(Oguma *et al.,* 1980; Oguma *et al.,* 1981). Samples of native BoNT/C, native BoNT/D, and rabbit anti-sera against these toxins (anti-C_{ST} and anti-D_{SA}) were kindly provided by C. Ellis (Applied Biotechnology, Onderstepoort Veterinary Institute, South Africa).

3.2.2 Construction and characterization of recombinant donor plasmids

3.2.2.1 Restriction enzyme digestions

The pFASTBAC™DUAL vector was digested with appropriate restriction enzymes to enable the cloning of the respective neurotoxin genes. To enable the cloning of the BoNT/D_{SA} and BoNT/MutD_{SA} genes, the pFASTBAC™DUAL vector was digested with both *EcoRI* and *SalI.* The pFASTBAC™DUAL vector was digested with *Eco*RI (Roche Diagnostics) in buffer H at 37°C for 30 min before *SalI* (Roche Diagnostics) was added and incubated for a further 45 min at 37°C. To enable the cloning of the BoNT/C_{ST} and BoNT/MutC_{ST} genes, the vector was digested with *SalI* (Roche Diagnostics) in buffer H

at 37°C for 1 h. The BoNT/C_{ST} and BoNT/MutC_{ST} genes were obtained from pGEM-C_{ST} and pGEMut-C_{ST} by *Sal*I restriction enzyme digestion. The BoNT/D_{SA} and BoNT/MutD_{SA} neurotoxin genes were obtained from pGEM-D_{SA} and pGEMut-D_{SA} by *Eco*RI and *Sal*I digestion.

3.2.2.2 Purification of DNA fragments from agarose gels

The restriction enzyme-digested BoNT genes and pFASTBAC™DUAL vectors were purified from a 0.8% (w/v) agarose gel using a silica suspension as described in section 2.2.6. The concentration of the purified DNA fragments was determined on a 0.8% (w/v) agarose gel as described in section 2.2.4.

3.2.2.3 Ligation of DNA fragments

The purified BoNT genes were cloned into the pFASTBAC™DUAL vector to construct donor plasmids. Ligation of the purified BoNT genes and pFASTBAC™DUAL vector was performed overnight (16 h) at 15°C in a total volume of 10 µl. The reaction mixture contained 1 µl of a 10x T4 DNA ligase buffer (660 mM Tris-HCl; 10 mM DTT; 50 mM $MgCl_2$; 10 mM ATP; pH 7.5), 50 ng of restricted pFastBac™Dual vector, ca. 200 ng insert DNA, 3 U of T4 DNA ligase (Roche Diagnostics, 3 U/µl), and UHQ water. The vector to insert ratio was typically 1:3.

3.2.2.4 Preparation of competent cells

Competent *E. coli* cells were prepared according to the Calcium/Manganese-based (CCMB) method as described by Hanahan *et al.* (1991). This method allows for the long-term storage of the competent cells. A bacterial suspension was prepared by inoculating a colony from a freshly-streaked *E. coli* JM109 plate into 2 ml SOB-Mg (20% (w/v) tryptone; 5% (w/v) yeast extract, 10 mM NaCl; 2.5 mM KCl; pH 7.0). The bacterial suspension was inoculated into 50 ml pre-heated SOB-Mg medium and incubated for 3 h at 30°C until the cells reached an OD_{550} value of 0.35 (mid-log phase). After incubation of the cells on ice for 10 min, cells from 30 ml of culture were collected in a Corex tube by centrifugation at 2 000 rpm for 10 min at 4°C in a pre-cooled Sorvall HB4 rotor. The cell pellet was dispersed in 10 ml pre-cooled CCMB80 (10 mM Potassium acetate; 10% (w/v) glyserol; 80 mM $CaCl_2$; 20 mM $MnCl_2$; 10 mM $MgCl_2$; pH 6.4) by gentle vortexing, incubated on ice for 20 min and pelleted as above. The pellet was

finally suspended in 2.5 ml pre-cooled CCMB80. The cells were aliquoted into sterile Eppendorf tubes and frozen in liquid nitrogen. The cells were stored at –70°C until use.

3.2.2.5 Transformation of competent cells

The competent *E. coli* JM109 cells were transformed after the addition of the ligation mixture using the heat shock method. The competent cells (100 µl) and the ligation mixture (5 µl) were mixed in a glass transformation tube and maintained on ice for 30 min. The cells were heat-shocked for 90 s at 42°C and chilled on ice for 2 min before pre-warmed SOC medium (900 µl) was added. The transformation mixtures were incubated with agitation at 37°C for 1 h to allow the cells to recover and to express the ampicillin and gentamycin resistance genes. The transformed cells were selected by plating the cells onto LB plates containing antibiotics (100 µg/ml ampicillin; 7 µg/ml gentamycin) and incubating overnight at 37°C.

pUC18 plasmid DNA (10 ng) was included as a transformation control and to calculate the competency of the cells. A negative control comprising only competent cells was also included to determine if any contamination occurred.

3.2.2.6 Screening of recombinant clones

The pFASTBAC™DUAL vector does not contain a selective marker to screen for recombinants. Recombinant clones were selected according to size with a rapid one-step screening procedure (Beuken *et al.*, 1998). Several colonies were inoculated into 3 ml LB broth containing antibiotics (100 µg/ml ampicillin; 7 µg/ml gentamycin) and incubated overnight at 37°C with shaking. Three hundred microliters (300 µl) of the overnight culture was centrifuged for 15-20 s after which the supernatant was discarded. The cells were lysed by briefly vortexing the pellet in 40 µl of loading dye (40% (w/v) sucrose; 0.25% (w/v) bromophenol blue) and 14 µl of a phenol:chloroform (1:1) mixture. The cell debris was collected by centrifugation at 13 000 rpm for 3 min. The samples were analyzed by loading 10 µl of the aqueous phase onto a 0.8% (w/v) agarose gel. Plasmid DNA migrating slower than the parental vector was selected for further characterization.

Plasmid DNA of putative recombinant clones was isolated by the alkaline lysis method as described by Birnbiom and Doly (1979) with slight modifications (discussed in section 2.2.7.1). DNA analysis of recombinant clones were performed on a 0.8% (w/v) agarose gel stained with EtBr for visualization on a transilluminator.

Recombinant plasmid DNA was digested with *EcoRI* (Roche Diagnostics) and *SalI* (Roche Diagnostics) to determine the presence and size of cloned DNA inserts. The reaction mixtures (15 – 20 μl) containing plasmid DNA (ca. 1 μg), 5 U of enzyme and the appropriate restriction buffer were incubated at 37°C for 45 min. The digestion products were analyzed on a 0.8% (w/v) agarose gel as described in section 2.2.4. Recombinant plasmid DNA was also digested with *BamHI* (Promega) and *EcoRI* (Promega) to determine the orientation of the cloned DNA inserts.

3.2.3 Construction of recombinant Bacmid DNA

3.2.3.1 Preparation of competent *E. coli* DH10Bac™ cells

Competent *E. coli* DH10Bac™ cells were prepared by the dimethyl sulfoxide (DMSO) method, as described by Chung and Miller (1988). An overnight culture was prepared by inoculating *E. coli* DH10Bac™ cells in 5 ml LB broth containing antibiotics (10 μg/ml tetracycline; 50 μg/ml kanamycin) with overnight shaking at 37°C. The overnight culture was inoculated into 100 ml of the same medium and grown to an OD_{600} value of 0.3 to 0.6 (early logarithmic phase). The cells were collected by centrifugation at 4 000 rpm for 5 min at 4°C. After centrifugation the cells were resuspended in 1/10th volume of ice-cold TSB (1.6% (w/v) peptone; 1% (w/v) yeast extract; 500 mM NaCl; 10% (w/v) PEG, MW 3350; 5% (v/v) DMSO; 10 mM $MgCl_2$; 10 mM $MgSO_4$) and incubated on ice for 20 min before transformation. Cells were also frozen at this stage and stored at –70°C for future use.

3.2.3.2 Transformation of competent *E. coli* DH10Bac™ cells

The donor plasmids (pFASTBACDUAL-C_{ST}, pFASTBACDUAL-MutC_{ST}, pFASTBACDUAL-D_{SA}, and pFASTBACDUAL-MutD_{SA})(500 ng) and the competent DH10Bac™ cells (100 μl) were gently mixed in glass transformation tubes and maintained on ice for 30 min. The cells were then heat-shocked for 45 s at 42°C and chilled on ice for 2 min before 900 μl of

pre-warmed SOC medium (2% (w/v) tryptone; 0.5% (w/v) yeast extract; 10 mM NaCl; 2.5 mM KCl; 10 mM $MgCl_2$; 20 mM glucose; pH 7.0) was added. The transformation mixtures were incubated with agitation at 37°C for 4 h to allow the cells to recover, to express the antibiotic resistance genes, and to allow for transposition. The transformed cells were selected by plating the cells onto Luria agar (1% (w/v) peptone; 0.5% (w/v) yeast extract; 0.17 M NaCl; pH 7.4) containing antibiotics (10 μg/ml tetracycline; 7 μg/ml gentamycin; 50 μg/ml kanamycin) in the presence of 10 μl IPTG (100 mM stock solution), and 40 μl X-gal (2% (w/v) stock solution). Plates were incubated for at least 24 h at 37°C. White colonies were selected and re-streaked onto fresh plates containing IPTG and X-gal to verify the phenotype.

Parental pFASTBAC™DUAL (500 ng) was included as a transformation control. A negative control comprising only competent cells was also included to determine if any contamination occurred.

3.2.3.3 Isolation of recombinant Bacmid DNA

The recombinant bacmid DNA was isolated by the following protocol that was developed for isolating large plasmids and adapted for isolating bacmid DNA. Three colonies confirmed as having a white phenotype on plates with IPTG and X-gal were grown in 2 ml of LB-medium containing antibiotics (50 μg/ml kanamycin; 7 μg/ml gentamycin; 10 μg/ml tetracycline) at 37°C with shaking for 24 h. After incubation, cells from 1.5 ml of each culture was collected in Eppendorf tubes by centrifugation at 13 000 rpm for 1 min. The supernatant was discarded and the bacterial pellet suspended in 300 μl of Solution I (15 mM Tris-HCl, pH 8.0; 10 mM EDTA; 100 μg/ml RNAse A). Three hundred microliters (300 μl) of Solution II (0.2 N NaOH; 1% (w/v) SDS) was then added before the tubes were incubated at room temperature for 5 min. Following the very slow addition of 300 μl of 3 M potassium acetate (pH 5.5) the tubes were incubated on ice for 10 min. The white precipitate consisting of protein and *E. coli* genomic DNA was removed by centrifugation at 13 000 rpm for 10 min. The bacmid DNA was precipitated from the supernatant by the addition of 800 μl isopropanol on ice for 5 to 10 min. Precipitated bacmid DNA was collected by centrifugation at 13 000 rpm for 15 min in a Hettich centrifuge. The residual salt was removed by washing the DNA pellet with 500 μl of 70% ethanol. The pellet was air-dried for 10 min at room temperature and suspended

in 40 µl of 1x TE buffer before storage at –20°C. The DNA was analyzed on a 0.8% (w/v) agarose gel as described in section 2.2.4.

3.2.4 Transfection of *S. frugiperda* cells

Spodoptera frugiperda (Sf9) insect cells were used for the transfection and propagation of the wild-type and recombinant baculoviruses using Graces insect medium (Highveld Biological (Pty) Ltd) or SFII-900 (Life Technologies) containing 5% FCS. Cell concentrations of viable cells were determined using a haemacytometer and staining with trypan blue.

Recombinant bacmid DNA was used to transfect *S. frugiperda* cells in 35 mm wells (1×10^6 cells/well) of a tissue culture dish. The cells were allowed to attach at 27°C for 1 h before the addition of the transfection solution. For each transfection a transfection solution was prepared by combining solution A (5 µl of bacmid DNA in 100 µl serum free medium (SFM) without antibiotics) and solution B (6 µl CELLFECTIN™ reagent in 100 µl SFM without antibiotics) by gentle mixing. The solutions were incubated at room temperature for 45 min before the addition of 800 µl of SFM to the lipid-DNA complexes. The attached Sf cells were washed with 2 ml of SFM before being overlaid with 1 ml of the diluted lipid-DNA complexes. The cells were incubated for 5 h at 27°C before the transfection mixtures were removed and replaced with 2 ml SFM containing antibiotics (Penicillin G sodium; streptomycin sulphate 10 mg/ml each; Fungizone 25 µg/ml in DRO water) (Highveld Biological (Pty) Ltd). The dishes were incubated at 27°C for 96 h. Virus stocks were harvested by centrifugation of cell suspensions at 2 000 rpm for 10 min in a Sorval centrifuge and the supernatants were stored at 4°C.

3.2.5 Titration and purification of baculovirus stocks

To obtain purified single plaques or to determine virus titer, plaque assays were performed. The Sf9 cells were seeded in 35 mm (1×10^6 cells/well) wells of a tissue culture dishes and allowed to attach at 27°C for 1 h. Serial dilutions of virus stocks (10^{-1} to 10^{-7}) were added to the Sf9 monolayers and incubated for 1 h to allow virus particles to infect the cells. The virus dilutions were removed and the cells were then overlaid

with an agarose-medium mixture (1.5% (w/v) low melting agarose in Graces medium or 1% (w/v) low melting agarose in SFII-900 medium) to keep cells stable and to limit the spread of virus. Dishes were incubated at 27°C for 4 days before the plaques were stained with 1 ml neutral red (1 mg/ml UHQ diluted 10x with medium) for 5 hours at 27°C. The Neutral red solution was subsequently removed and the dishes were incubated overnight. Single plaques were picked and subjected to three rounds of amplification before the titer of the viral stocks was determined. The recombinant baculoviruses containing the wild-type BoNT/C gene (Baculo-BoNT/C), mutant BoNT/C gene (Baculo-BoNT/MC), wild-type BoNT/D gene (Baculo-BoNT/D), and mutant BoNT/D gene (Baculo-BoNT/MD) were stored at 4°C after filter-sterilization of the viral stocks.

3.2.6 Analysis of recombinant baculoviruses

3.2.6.1 Bacmid DNA preparation

Bacmid DNA for PCR analysis was prepared by adding 100 μl of the virus stock to 890 μl lysis buffer (10 mM Tris-HCl, pH 8.3; 100 μg/ml gelatin; 0.45% (v/v) Triton X-100; 50 mM KCl) containing 6 μg Proteinase K. The samples were incubated at 60°C for 1 h before the proteinase K was heat inactivated by boiling the samples at 95°C for 10 min. The bacmid DNA was subsequently precipitated at –70°C for 2 h following the addition of 2 volumes 96% ethanol and sodium acetate (pH 7.0) to a final concentration of 0.3 M. Precipitated DNA was pelleted by centrifugation at 13 000 rpm for 15 min, washed twice with 70% ethanol, dried under vacuum and resuspended in 30 μl 1x TE buffer (10 mM Tris-HCl, pH 7.4; 1 mM EDTA, pH 8.0).

3.2.6.2 Selection of primers

Oligonucleotide primers BNTseqC4, BNTseqD4, BNTseq18, Cmut-F, and Dmut-F (Table 2.1 and 2.2) were used in a heminested PCR in order to verify the presence of the wild-type and mutant BoNT/C and BoNT/D genes in the recombinant bacmid DNA. Primers were designed according to published sequences of the type C and D genes (Kimura *et al.*, 1990; Moriishi *et al.*, 1996a). The primers used for heminested PCR on the wild-type and mutant BoNT/C gene were designed to anneal at the following positions: BNTseqC4 at nt 419 – 440, Cmut-F at nt 685 – 715, and BNTseq18 at nt 1858 – 1880. The primers used for heminested PCR on the wild-type and mutant BoNT/D genes were designed to

anneal at the following positions: BNTseqD4 at nt 391 – 412, Dmut-F at nt 685 – 714, and BNTseq18 at nt 1846 – 1868.

3.2.6.3 Heminested PCR

The reaction mixtures (50 µl) used for both PCR steps contained 5 µl of 10x Taq DNA polymerase buffer (500 mM KCl; 100 mM Tris-HCl, pH 9.0; 1.0% Triton X-100), 1.5 mM $MgCl_2$, 200 µM of each dNTP, 1 U of Taq DNA polymerase (Promega), and ultra high quality (UHQ) water. For the first PCR step, 10 ρmol of each of the sense (BNTseqC4 or BNTseqD4) and anti-sense (BNTseq18) primers, and 1 µl of recombinant bacmid DNA were used. The reaction tubes were placed in a Perkin-Elmer 2400 thermal cycler (Perkin-Elmer, Cetus, Norwalk, CT, USA). The samples were subjected to 30 of the following cycles: denaturation at 94°C for 15 s, primer annealing at 40°C for 30 s, and elongation at 72°C for 2 min. After the last cycle, the samples were kept at 72°C for 6 min to complete synthesis of all strands. The second PCR step was performed using 10 ρmol of each of the sense (Cmut-F or Dmut-F) and anti-sense (BNTseq18) primers, and 1 µl of the first PCR product (diluted 1:10) as template. The cycle profile consisted of the same heat denaturation, primer annealing, and DNA extension conditions as those used for the first PCR step. A positive control consisting of plasmid DNA containing the respective genes was included during each amplification step. A control containing wild-type bacmid DNA was included to test for non-specific amplification and a negative control in which DNA was omitted was also included. The amplicons were analyzed on a 1% (w/v) agarose gel as described in section 2.2.4

3.2.7 Fractionation of infected *S. frugiperda* cells

S. frugiperda cells in 75 cm^2 tissue culture flasks (6 x 10^9 cells/flask) were infected with either Baculo-BoNT/C, Baculo-BoNT/MC, Baculo-BoNT/D, Baculo-BoNT/MD, or wild-type baculovirus at a MOI of 5 pfu/cell. The infected monolayers were incubated at 27°C for 72 h before the cells were harvested, collected by centrifugation at 3 000 rpm for 10 min in a Beckman Ultra Centrifuge and washed twice with 1x PBS. The cells were resuspended in 600 µl 10 mM STE buffer (10 mM NaCl; 1 mM EDTA; 10 mM Tris-HCl, pH 7.4) and incubated on ice for 10 min. Following homogenization of the cell suspension by 15 strokes with a Dounce homogenizer, the nuclei were removed by

centrifugation at 2 000 rpm for 2 min. The supernatants (S10 fractions) were collected before the addition of three protease inhibitors (0.5 mM Phenylmethylsulfonyl Fluoride (PMSF); 1 μg/ml Leupeptin; 5 μg/ml Aprotinin) to limit protease activity. The S10 fractions were stored at –20°C.

3.2.8 SDS-PAGE analysis of proteins synthesized in infected *S. frugiperda* cells

The proteins synthesized in infected *S. frugiperda* cells were subjected to separation according to size by sodium dodecyl sulfate polyacrylamide-gel electrophoresis (SDS-PAGE). SDS-PAGE analysis was carried out by using a discontinuous buffer system as described by Laemmli (1970). A bipartite gel was used, comprising a high porosity stacking gel (5% (w/v) acrylamide; 0.17% (w/v) bisacrylamide; 125 mM Tris-HCl, pH 6.8; 0.1% (w/v) SDS; ammonium persulfate; tetramethylethylenediamine (TEMED)) and a low porosity separating gel (10% (w/v) acrylamide; 0.34% (w/v) bisacrylamide; 0.375 M Tris-HCl, pH 8.7; 0.1% (w/v) SDS; ammonium persulfate; TEMED). The proteins were suspended in 2x PSB (125 mM Tris-HCl, pH 6.8; 4% (w/v) SDS; 20% (v/v) glycerol; 0.002% (w/v) bromophenol blue; 10% (v/v) 2-mercaptoethanol) and boiled for 5 min before being loaded onto the gel. The gel was electrophoresed at 125 V for 2 to 3 h in 1x TGS buffer (25 mM Tris; 192 mM glysine, 1% (w/v) SDS). The gels were then stained with Coomassie brilliant blue (0.125% (w/v) Coomassie; 50% (v/v) methanol; 10% (v/v) acetic acid) for 5 min and subsequently destained with destaining solution (10% (v/v) acetic acid; 10% (v/v) methanol). The protein sizes were determined by comparison to a high-range molecular weight marker (Amersham Pharmacia Biotech).

3.2.9 Immunological detection of expressed proteins

3.2.9.1 Preparation of acetone powders and adsorption of anti-sera

Acetone powders of uninfected *S. frugiperda* cells were prepared as described by Sambrook *et al.* (1989). *S. frugiperda* cells from suspension cultures were collected by centrifugation at 6 000 rpm for 10 min. The cell pellet was suspended in ice-cold 0.1 M NaCl (at 2 ml/g wet weight) and homogenized by 15 strokes with a Dounce homogenizer. Four volumes of acetone at –20°C were added before incubation on ice for 1 h. The precipitate was collected by centrifugation at 7 500 rpm for 10 min at 4°C, resuspended in fresh acetone at –20°C, incubated on ice for 10 min and again collected

by centrifugation. The pellet was lyophilized and stored air-tight at –20°C. Adsorption of the anti-sera against the acetone powder was performed by adding 15 mg acetone powder to 100 µl of the anti-sera (anti-C_{ST}(L) and anti-D_{1873}) and incubation on ice for 30 min. After centrifugation at 7 500 rpm for 10 min at 4°C, the supernatant was retained and stored at –20°C until use.

3.2.9.2 Western blot analysis

A duplicate SDS-PAGE gel was subjected to wet blotting onto a Hybond™-C super nitrocellulose membrane for Western blot analysis. The gel and nitrocellulose membrane, cut to the same size as the gel, were soaked in transfer buffer (40 mM glysine; 120 mM Tris) for 10 min before the proteins were transferred onto the membrane at 28 V, 120 mA for 2 h in a TE series Transphor Electrophoresis Unit (Hoefer Scientific Instruments) containing transfer buffer. Following transfer, the gels were recovered and stained with Coomassie Blue to determine the efficiency of the transfer process. The membranes were blocked by immersion in 1x PBS (13.7 mM NaCl; 0.27 mM KCl; 0.43 mM Na_2HPO_4; 0.14 mM KH_2PO_4, pH 7.3) containing 1% (w/v) fat-free milk powder for 16 h at 4ºC. Different anti-sera were diluted in fresh blocking solution before their addition to the respective membranes. Anti-C_{ST}(L) was diluted 1:30, pre-adsorbed anti-C_{ST}(L) was diluted 1:20, anti-D_{1873} was diluted 1:40, and pre-adsorbed anti-D_{1873} was diluted 1:20. Anti-C_{ST} and anti-D_{SA} were diluted 1:100 in fresh blocking solution. The membranes were gently agitated at room temperature for 2 h after which the unbound primary antibodies were removed by washing the membranes three times for 5 min each in wash buffer (0.05% (v/v) Tween-20 in 1x PBS). The secondary antibody (Protein A conjugated with horseradish peroxidase, diluted 1:1000 in fresh blocking solution)(Sigma Aldrich) was added to the membrane and incubated with agitation at room temperature for 1 h. The membranes were thoroughly washed as before with wash buffer (0.05% (v/v) Tween-20 in 1x PBS) before visualization of the antigen-antibody complex with peroxidase substrate (60 mg 4-chloro-1-naftol in 20 ml methanol; 60 µl hydrogen peroxide in 100 ml 1x PBS). The membranes were incubated in the substrate until the bands became visible. Positive controls consisting of purified BoNT/C and BoNT/D were included in the respective Western blots.

3.3 RESULTS

3.3.1 Construction of donor plasmids containing the wild-type and mutant type C and D neurotoxin genes

Full-length copies of the respective neurotoxin genes were obtained by restriction enzyme digestion of recombinant pGEM®-T Easy vectors previously constructed as described in section 2.2. The BoNT/C and BoNT/MC genes were recovered with *SalI* restriction enzyme digestion from pGEM-C_{ST} and pGEMut-C_{ST} for cloning into *SalI*-digested pFASTBAC™DUAL vector (Fig. 3.1). The BoNT/D and BoNT/MD genes were recovered from pGEM-D_{SA} and pGEMut-D_{SA} by digestion with both *EcoRI* and *SalI* to allow for directional cloning into linearized pFASTBAC™DUAL vector (Fig. 3.1).

The gel-purified DNA fragments were subsequently ligated with linearized pFASTBAC™DUAL vector (Life Technologies). Recombinant plasmids were selected according to size following a rapid one-step screening procedure. The isolated plasmid DNA was characterized by agarose and restriction enzyme analysis. Plasmid DNA migrating slower than the parental pFASTBAC™DUAL vector were selected and digested with *EcoRI* and *SalI* to determine whether the plasmid DNA contained cloned inserts. *SalI* digestion of recombinant pFASTBACDUAL-C_{ST} and pFASTBACDUAL-MutC_{ST} resulted in two fragments of ca. 3.8 kb and 5.2 kb, respectively (Fig. 3.2). This corresponds with the expected size of the DNA insert and parental linearized vector. The recombinant pFASTBACDUAL-C_{ST} and pFASTBACDUAL-MutC_{ST} plasmids were also digested with *EcoRI* to determine the orientation of the cloned inserts. *EcoRI* digestion resulted in two fragments of ca. 1.6 kb and 7.4 kb, respectively (Fig. 3.2). This corresponds with the expected sizes according to partial restriction enzyme maps. The digestion of recombinant pFASTBACDUAL-D_{SA} and pFASTBACDUAL-MutD_{SA} with both *EcoRI* and *SalI* also resulted in two fragments of ca. 3.8 kb and 5.2 kb, respectively (Fig. 3.4). This corresponds with the expected size of the DNA insert and parental linearized vector (Fig. 3.3). The recombinant pFASTBACDUAL-D_{SA} and pFASTBACDUAL-MutD_{SA} plasmids were also digested with *BamHI* to verify the orientation of the directionally cloned inserts. *BamHI* digestion resulted in two fragments of ca. 660 bp and 8.4 kb, respectively (Fig. 3.4). This corresponds with the expected sizes according to partial restriction enzyme maps (Fig. 3.3).

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

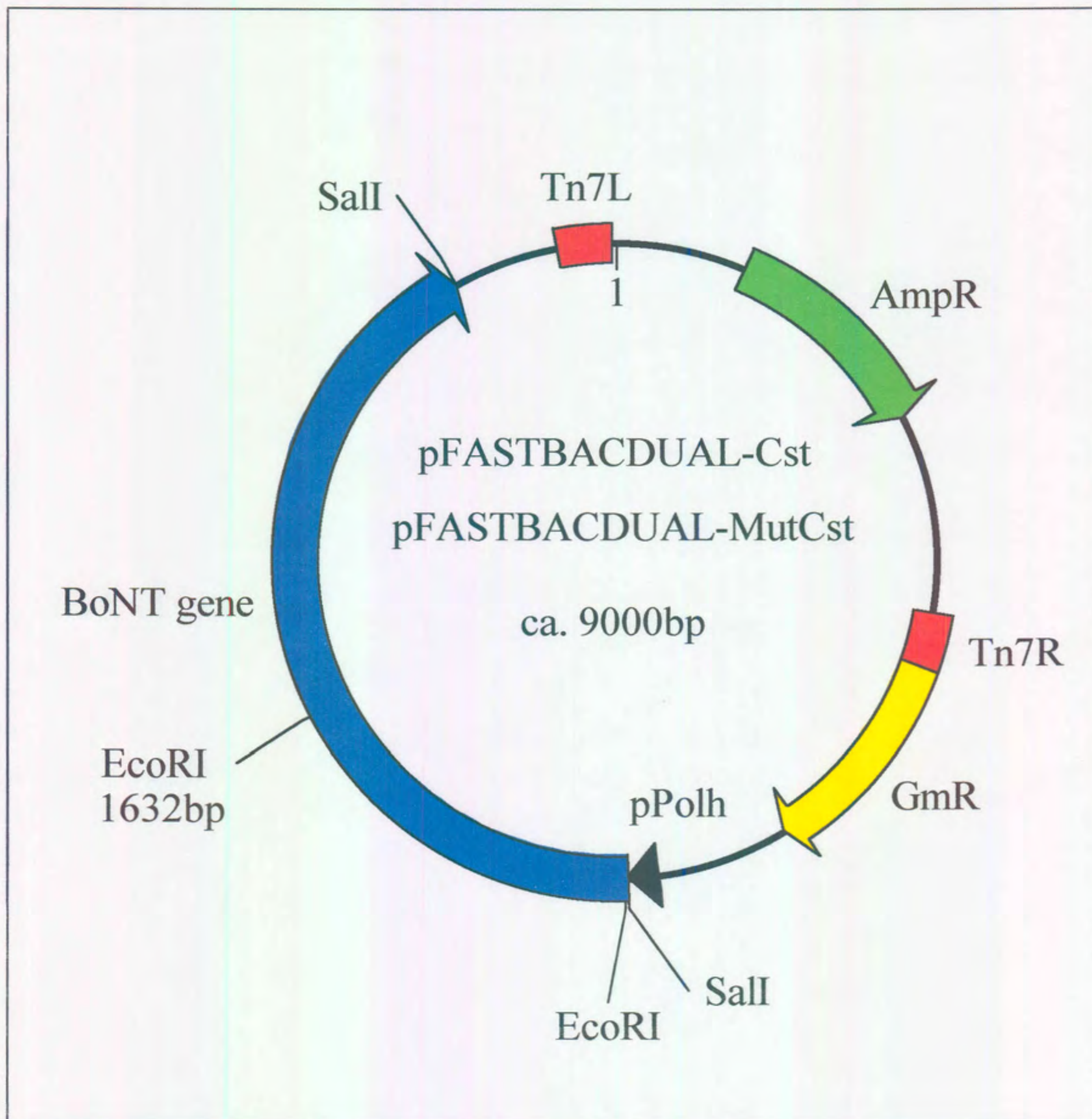

Fig. 3.1 Partial restriction enzyme map of recombinant pFASTBAC™DUAL vector containing either the BoNT/C_{ST} or BoNT/MutC_{ST} gene.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

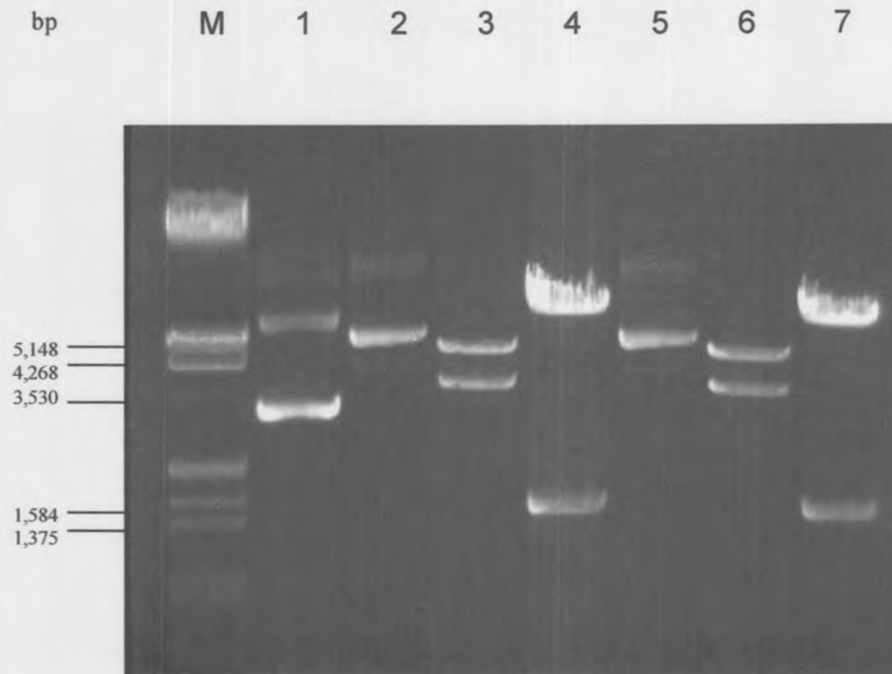

Fig. 3.2 Agarose gel electrophoretic analysis of the recombinant donor plasmids, pFASTBACDUAL-C_{ST} and pFASTBACDUAL-MutC_{ST}, constructed by cloning of the respective BoNT DNA fragments into linearized pFASTBAC™DUAL vector. Lanes: M, λ DNA/*EcoRI* + *HindIII* Marker; 1, Uncut non-recombinant pFASTBAC™DUAL vector; 2, Uncut vector pFASTBACDUAL-C_{ST}; 3, *SalI*-digested pFASTBACDUAL-C_{ST} vector; 4, *EcoRI*-digested pFASTBACDUAL-C_{ST} vector; 5, Uncut pFASTBACDUAL-MutC_{ST} vector; 6, *SalI*-digested pFASTBACDUAL-MutC_{ST} vector; 7, *EcoRI*-digested pFASTBACDUAL-MutC_{ST}

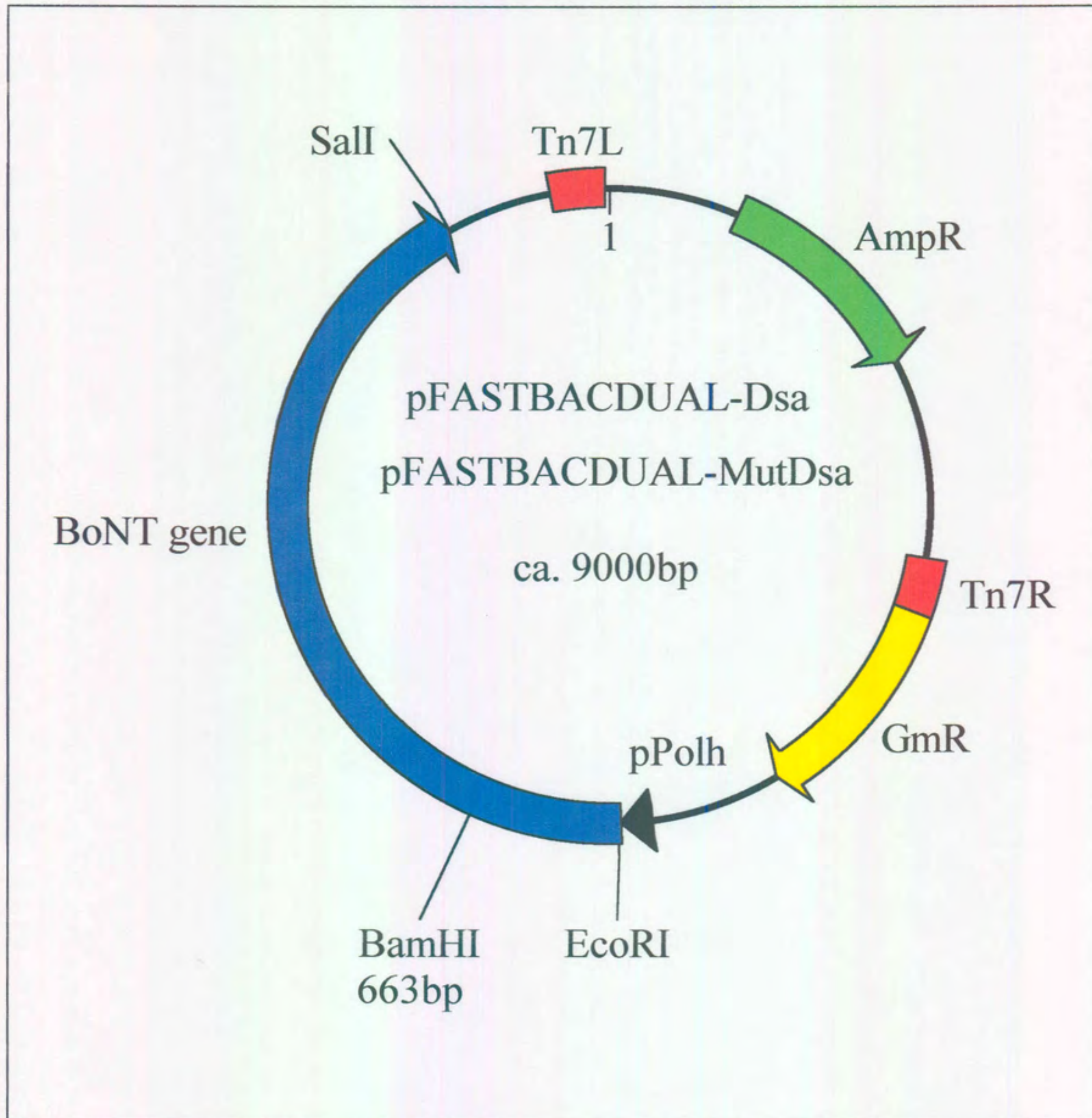

Fig. 3.3 Partial restriction enzyme map of recombinant pFASTBAC™DUAL vector containing either the BoNT/D_{SA} or BoNT/MutD_{SA} gene.

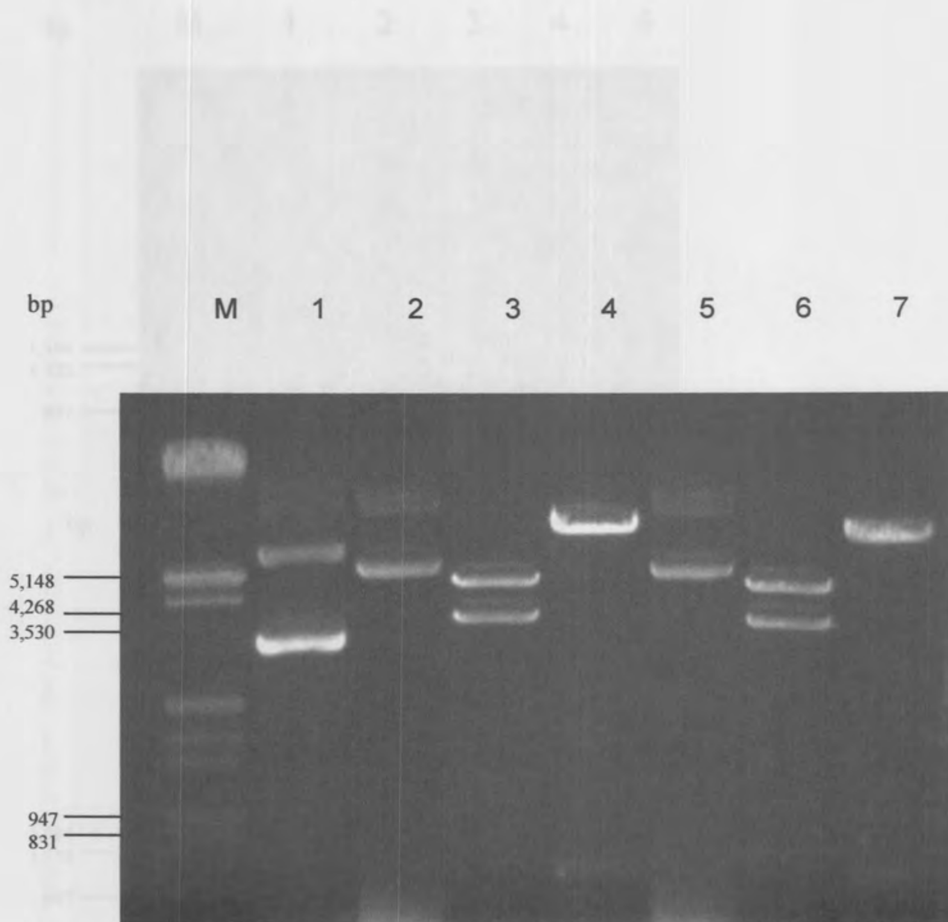

Fig. 3.4 Agarose gel electrophoretic analysis of the recombinant donor plasmids, pFASTBACDUAL-D_{SA}, and pFASTBACDUAL-MutD_{SA} constructed by cloning of the respective BoNT DNA fragments into linearized pFASTBAC™DUAL vector. Lanes: M, λ DNA/*EcoRI* + *HindIII* Marker; 1, Uncut non-recombinant pFASTBAC™DUAL vector; 2, Uncut vector pFASTBACDUAL-D_{SA}; 3, *SalI* and *EcoRI*-digested pFASTBACDUAL-D_{SA} vector; 4, *BamHI*-digested pFASTBACDUAL-D_{SA} vector; 5, Uncut pFASTBACDUAL-MutD_{SA} vector; 6, *SalI* and *EcoRI*-digested pFASTBACDUAL-MutD_{SA} vector; 7, *BamHI*-digested pFASTBACDUAL-MutD_{SA}.

3.3.2 Construction of recombinant baculoviruses

Recombinant plasmid DNA containing the respective wild-type and mutant neurotoxin genes were selected as donor plasmids for transposition to the target bacmid DNA in *E. coli* DH10Bac™. The bacmid DNA contained in *E. coli* DH10Bac™, propagates as a large plasmid that confers resistance to kanamycin and it can complement a *lacZ* deletion present on the *E. coli* chromosome to form colonies that are blue in the presence of X-gal and the inducer IPTG. Insertion of the mini-Tn7 from the donor plasmid into the mini-*att*Tn7 attachment site on the bacmid disrupts expression of the lacZα peptide, so colonies containing the recombinant bacmid are white in a background of blue colonies that harbor the unaltered bacmid. This allowed for the phenotypic selection of recombinant bacmid DNA.

Recombinant bacmid DNA was isolated from *E. coli* DH10Bac™ cells and transfected into insect cells using CELLFECTIN™ Reagent (Life Technologies). Following two rounds of amplification in monolayer cultures, recombinant viruses were plaque-purified and subjected to a further three rounds of amplification in monolayer cultures. The viral titers were determined by plaque assays and these were typically ca. 5.5×10^7 pfu/ml.

The presence of BoNT genes in the baculovirus genome was verified by a heminested PCR approach. A discreet band of 1.4 kb was observed for the first PCR step following agarose gel electrophoresis (Fig. 3.5(A) and Fig. 3.6(A)). This corresponds with the distance between the annealing sites of the sense primers (BNTseqC4 or BNTseqD4) and anti-sense primer (BNTseq18) on the type C and D genes, respectively. The obtained amplicons were diluted and used as template in the second PCR step. A discreet band of 1.2 kb was observed following agarose gel electrophoresis (Fig. 3.5(B) and Fig. 3.6(B)). This corresponds with the distance between the annealing sites of the internal sense primers (Cmut-F or Dmut-F) and anti-sense primer (BNTseq18) on the type C and D genes, respectively. In both rounds of PCR, amplicons of expected sizes were observed for the positive control consisting of pGEM®-T Easy vectors containing the respective genes. No amplification products were observed in the control in which wild-type bacmid DNA was used as well as the control in which template DNA was omitted.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

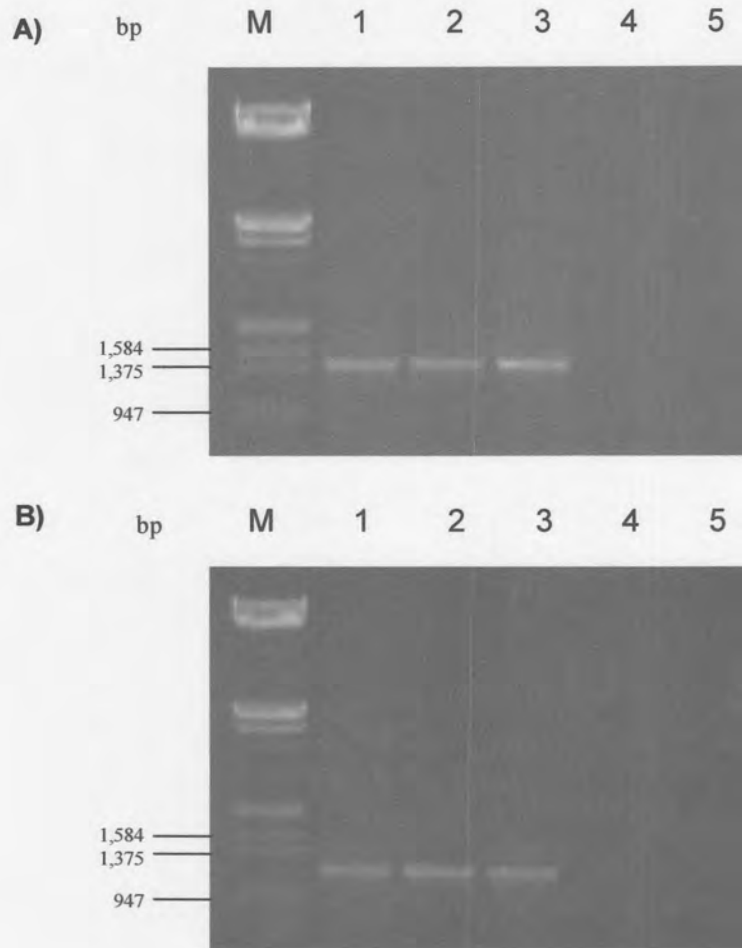

Fig. 3.5 Agarose gel electrophoresis of amplicons obtained by heminested PCR for verification of the presence of BoNT/C and BoNT/MC genes in the recombinant Bacmid DNA. The first PCR step (A) was performed with primer set 1 (BNTseqC4 and BNTseq18) using recombinant bacmid DNA as template. Lanes: M, λ DNA/*EcoRI* + *HindIII* Marker; 1, Baculo-BoNT/C bacmid DNA; 2, Baculo-BoNT/MC bacmid DNA; 3, pGEM-C_{ST} plasmid; 4, Wild-type Bacmid DNA control; 5, Water negative control. The second PCR step (B) was performed with primer set 2 (Cmut-F and BNTseq18) using the respective amplicons obtained in the first PCR step as template. Lanes: M, λ DNA/*EcoRI* + *HindIII* Marker; 1, Baculo-BoNT/C amplicon; 2, Baculo-BoNT/MC amplicon; 3, pGEM-C_{ST} amplicon; 4, Wild-type Bacmid DNA control; 5, Water negative control.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

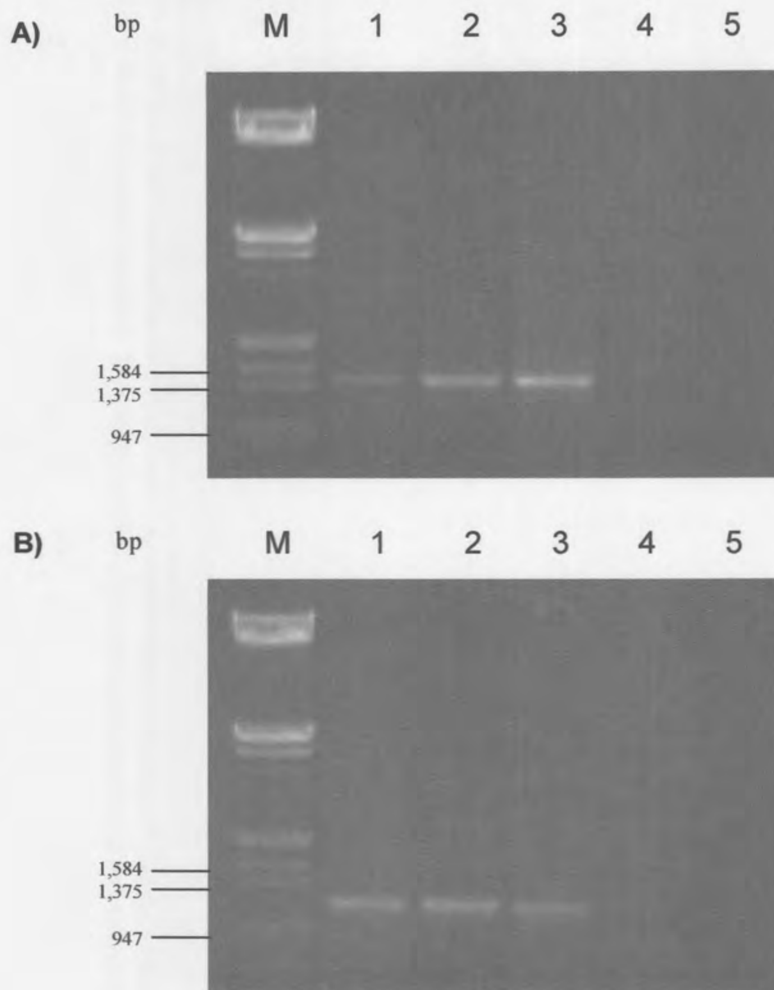

Fig. 3.6 Agarose gel electrophoresis of amplicons obtained by heminested PCR for verification of the presence of BoNT/D and BoNT/MD genes in the recombinant Bacmid DNA. The first PCR step (A) was performed with primer set 1 (BNTseqD4 and BNTseq18) using recombinant bacmid DNA as template. Lanes: M, λ DNA/*EcoRI* + *HindIII* Marker; 1, Baculo-BoNT/D bacmid DNA; 2, Baculo-BoNT/MD bacmid DNA; 3, pGEM-D_{SA} plasmid; 4, Wild-type Bacmid DNA control; 5, Water negative control. The second PCR step (B) was performed with primer set 2 (Dmut-F and BNTseq18) using the respective amplicons obtained in the first step as template. Lanes: M, λ DNA/*EcoRI* + *HindIII* Marker; 1, Baculo-BoNT/D amplicon; 2, Baculo-BoNT/MD amplicon; 3, pGEM-D_{SA} amplicon; 4, Wild-type Bacmid DNA control; 5, Water negative control.

3.3.3 Analysis of proteins synthesized in infected *S. frugiperda* cells

Cytoplasmic extracts of mock- and baculovirus-infected *S. frugiperda* cells were subjected to SDS-PAGE analysis under reducing conditions in the presence of 2-mercaptoethanol. Analysis of the Coomassie-stained gels (Fig. 3.7) indicated that no unique proteins could be observed in the recombinant baculovirus-infected cells compared to the mock- or wild-type infected cells.

Western blot analysis of cell extracts prepared from *S. frugiperda* cells infected with Baculo-BoNT/C or Baculo-BoNT/MC was performed using anti-C_{ST} and anti-C_{ST}(L) sera, respectively. The anti-C_{ST} serum was prepared against full-length BoNT/C_{ST}, while the anti-C_{ST}(L) serum was prepared against only the light chain of the type C_{ST} toxin (Oguma *et al.,* 1980). Western blot analysis performed using anti-C_{ST} serum indicated the presence of immunoreactive material migrating at positions of 150 kDa, 100 kDa, and 50 kDa in extracts prepared from wild-type baculovirus-infected *S. frugiperda* cells (Fig 3.8(A)). These non-specific bands migrated at the same positions as bands observed for the purified type C toxin (150 kDa) as well as the heavy (100 kDa) and light (50 kDa) chains of the reduced purified type C toxin, respectively. The immunoreactive material migrating at the position of the heavy chain was also detected in extracts prepared from *S. frugiperda* cells infected with Baculo-BoNT/C or Baculo-BoNT/MC. However, no specific immunoreactive material was detected and several non-specific background bands were observed.

Western blot analysis performed with anti-C_{ST}(L) serum indicated immunoreactive material migrating at a position of ca. 60 kDa in extracts prepared from *S. frugiperda* cells infected with Baculo-BoNT/C (Fig 3.8(B)). Immunoreactive material migrating at a position greater than 69 kDa was demonstrated for extracts prepared from *S. frugiperda* cells infected with Baculo-BoNT/MC. The immunoreactive band may correspond to the toxin light chain. The cell extracts prepared from the wild-type baculovirus-infected and mock-infected *S. frugiperda* cells did not exhibit these bands. However, several non-specific bands were observed. Anti-C_{ST}(L) also cross-reacted with the heavy chain of the reduced purified type C toxin. A non-specific immunoreactive band migrating at the same position as the heavy chain (100 kDa) was observed for all the cell extracts.

Western blot analysis of cell extracts prepared from *S. frugiperda* cells infected with Baculo-BoNT/D or Baculo-BoNT/MD was performed using both anti-D_{SA} and anti-D_{1873} sera. The anti-D_{SA} serum was prepared against full-length BoNT/D_{SA}, while the anti-D_{1873} serum was prepared against full-length BoNT/D_{1873} (Oguma *et al.*, 1981). Western blot analysis performed using anti-D_{SA} serum indicated immunoreactive material migrating at the same position as the heavy chain (100 kDa) in extracts prepared from *S. frugiperda* cells infected with Baculo-BoNT/D and Baculo-BoNT/MD (Fig. 3.9(A)). However, this band was also present in the mock- infected cells. No specific immunoreactive material was detected in extracts prepared from *S. frugiperda* cells infected with Baculo-BoNT/D or Baculo-BoNT/MD.

Western blot analysis performed using anti-D_{1873} serum indicated no specific immunoreactive material in extracts prepared from *S. frugiperda* cells infected with Baculo-BoNT/D or Baculo-BoNT/MD (Fig. 3.9(B)). Anti-D_{1873} serum cross-reacted with the light and heavy chains of the reduced purified type D_{SA} toxin. Several non-specific bands migrating at a position of between 55 kDa and 80 kDa were observed for all the samples. Pre-adsorped anti-sera were used in order to eliminate the non-specific background. However, results obtained with the pre-adsorbed anti-sera were the same as that for the normal anti-sera (results not shown).

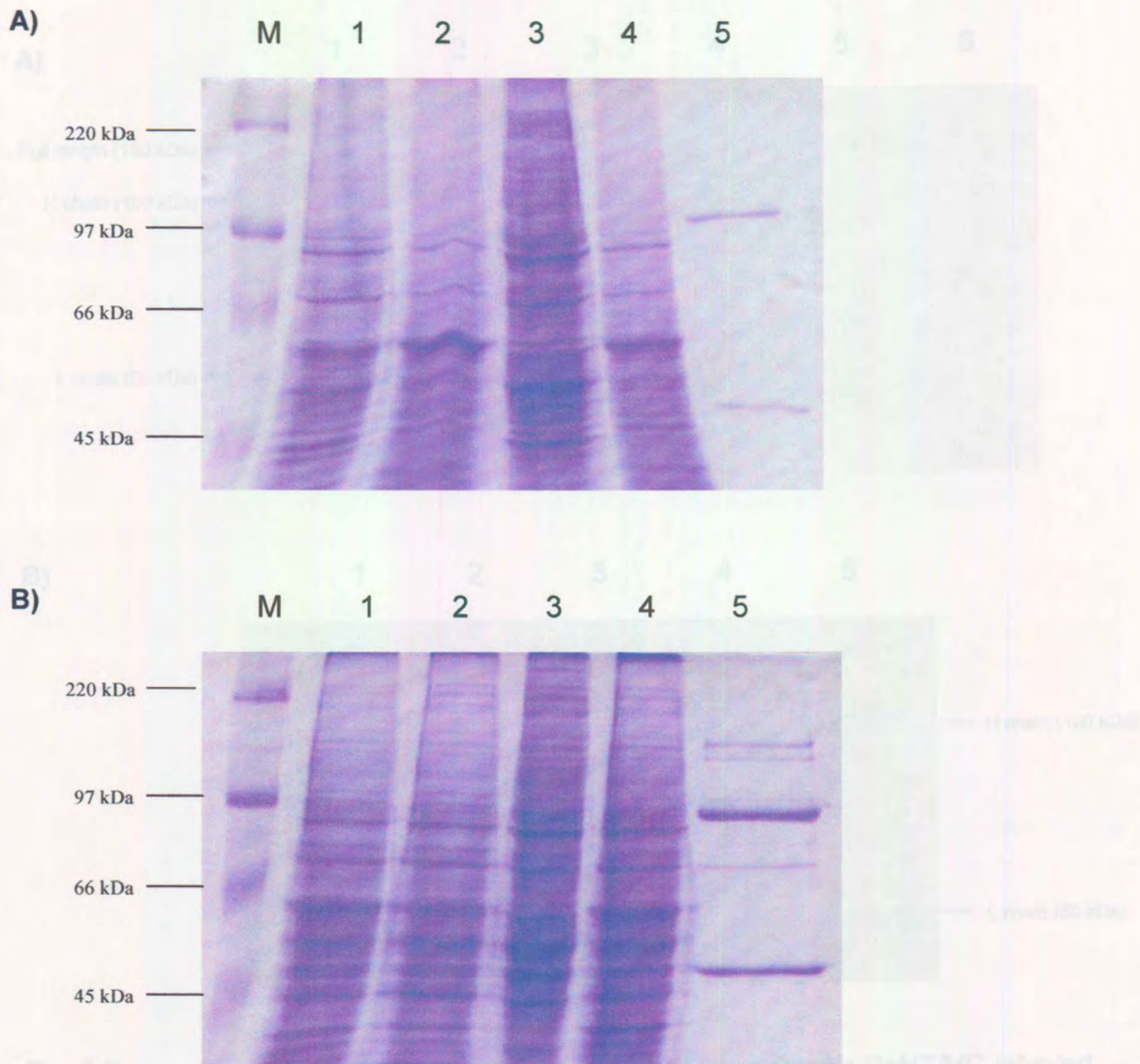

Fig. 3.7 SDS-polyacrylamide gel electrophoresis of recombinant baculovirus-infected *S. frugiperda* cells under reducing conditions in the presence of 2-mercaptoethanol. (A) Lanes: M, MW Marker; 1, Baculo-BoNT/C infected cells; 2, Baculo-BoNT/MC infected cells; 3, Mock infected cells; 4, Wild-type baculovirus infected cells; 5, Purified BoNT/C_{ST}. (B) Lanes: M, MW Marker; 1, Baculo-BoNT/D infected cells; 2, Baculo-BoNT/MD infected cells; 3, Mock infected cells; 4, Wild-type baculovirus infected cells; 5, Purified BoNT/D_{SA}.

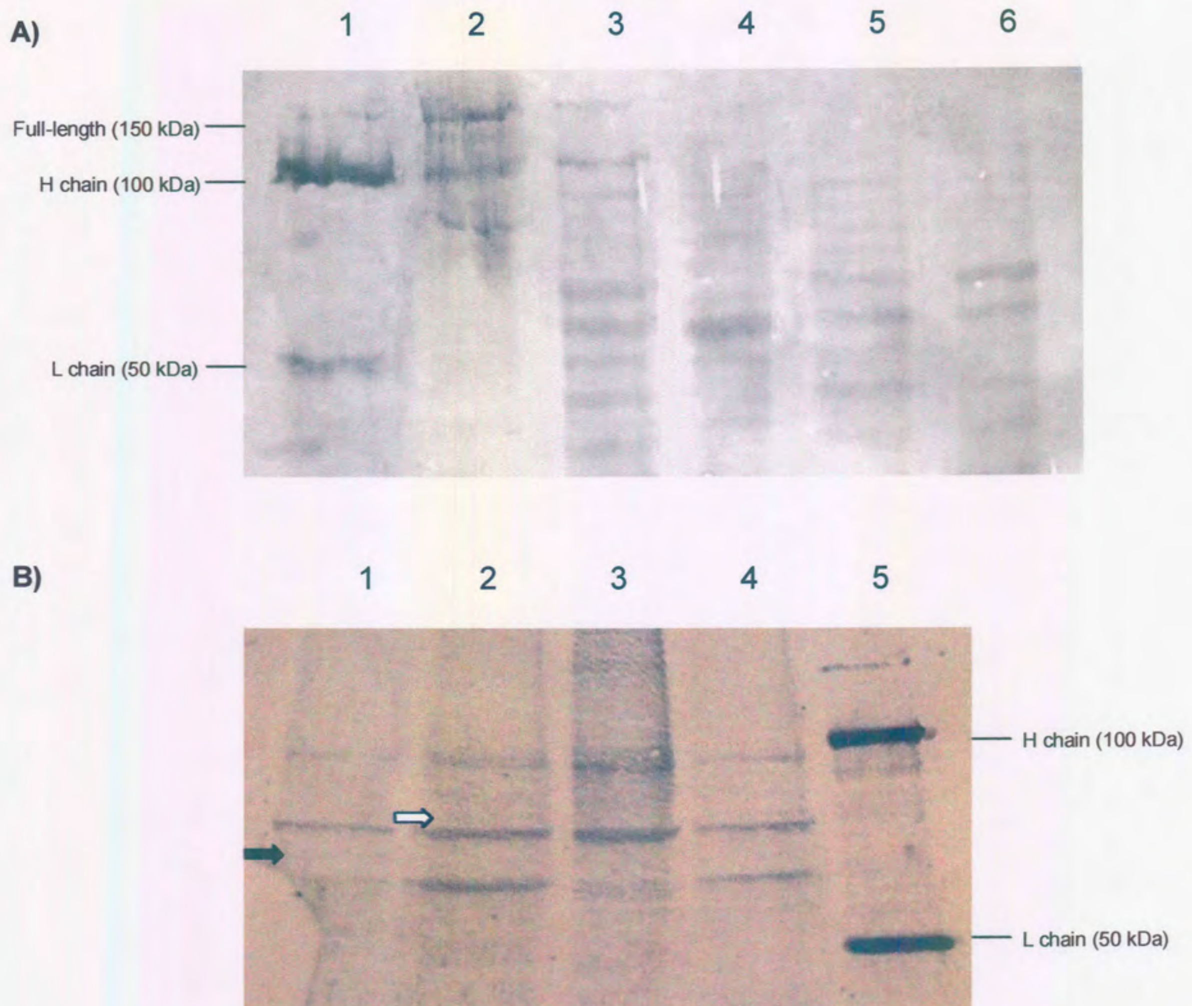

Fig. 3.8 Western blot analysis of Baculo-BoNT/C and Baculo-BoNT/MC infected *S. frugiperda* cell extracts subjected to SDS-PAGE electrophoresis under reducing conditions in the presence of 2-mercaptoethanol. (A) Analysis performed with anti-C_{ST}. Lanes: 1, Reduced purified BoNT/C_{ST}; 2, Non-reduced purified BoNT/C_{ST}; 3, Wild-type baculovirus infected cells; 4, Mock infected cells; 5, Baculo-BoNT/C infected cells; 6, Baculo-BoNT/MC infected cells. (B) Analysis performed with anti-C_{ST}(L). Lanes: 1, Baculo-BoNT/C infected cells; 2, Baculo-BoNT/MC infected cells; 3, Mock infected cells; 4, Wild-type baculovirus infected cells; 5, Reduced purified BoNT/C_{ST}. Specific immunoreactive material for Baculo-BoNT/C- and Baculo-BoNT/MC-infected cells are indicated by closed and opened arrows, respectively.

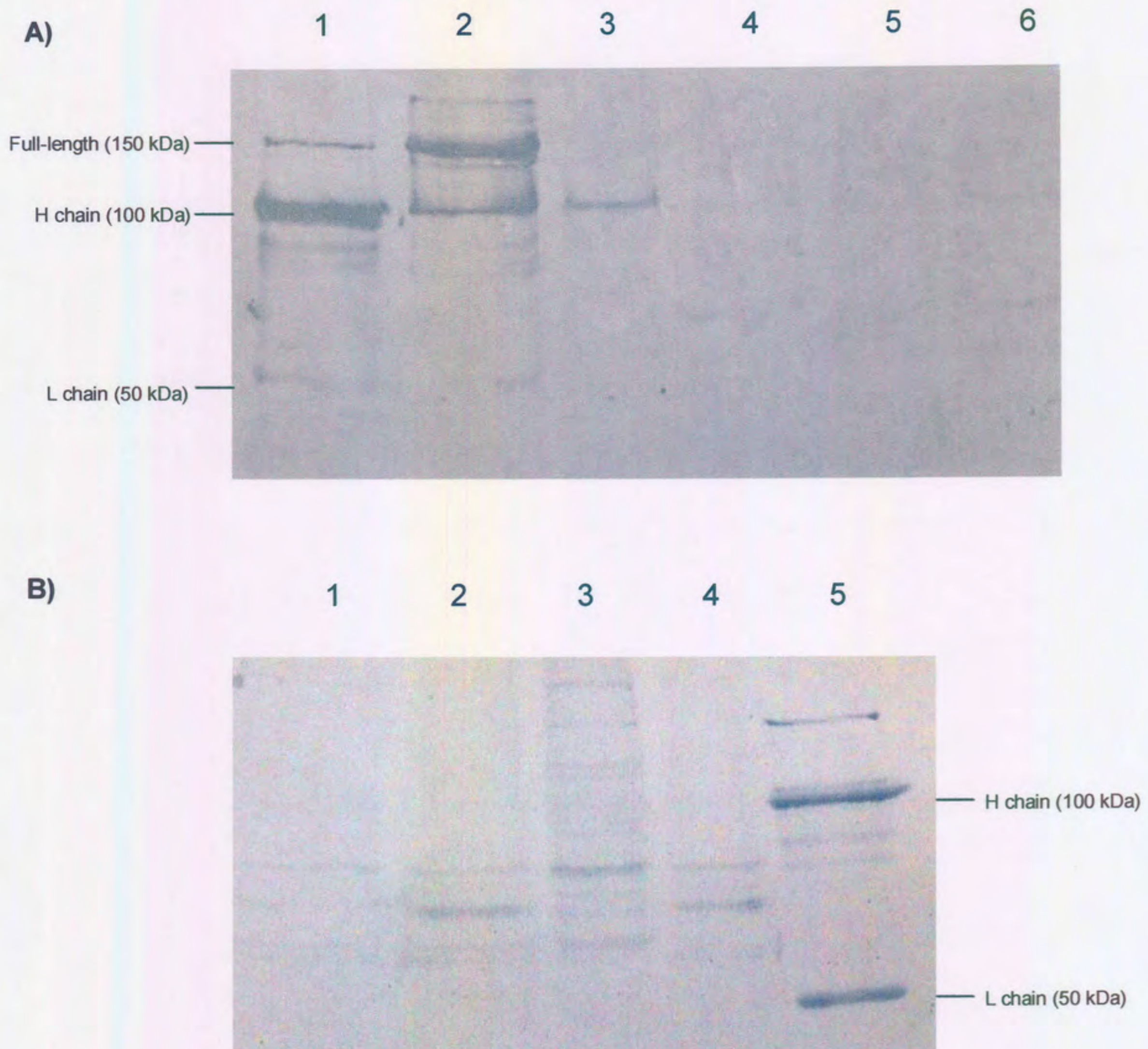

Fig. 3.9 Western blot analysis of Baculo-BoNT/D and Baculo-BoNT/MD infected *S. frugiperda* cell extracts subjected to SDS-PAGE electrophoresis under reducing conditions in the presence of 2-mercaptoethanol. (A) Analysis performed with anti-D_{SA}. Lanes: 1, Reduced purified BoNT/D_{SA}; 2, Non-reduced purified BoNT/D_{SA}; 3, Baculo-BoNT/MD infected cells; 4, Baculo-BoNT/D infected cells; 5, Wild-type baculovirus infected cells; 6, Mock infected cells. (B) Analysis performed with anti-D_{1873}. Lanes: 1, Baculo-BoNT/D infected cells; 2, Baculo-BoNT/MD infected cells; 3, Mock infected cells; 4, Wild-type baculovirus infected cells; 5, Reduced purified BoNT/D_{SA}.

3.4 DISCUSSION

The past decade has seen various attempts to express the botulinum neurotoxin genes in several different expression systems. The baculovirus expression system has gained widespread use for the production of a wide variety of heterologous proteins (O'Reilly *et al.,* 1992 and references therein). In this chapter the expression of the full-length wild-type and mutant BoNT/C_{ST} and BoNT/D_{SA} genes in *S. frugiperda* cells using the BAC-to-BAC™ baculovirus system was investigated for its usefulness as an alternative expression system for botulinum neurotoxin genes.

The wild-type and mutant full-length botulinum neurotoxin genes were cloned into the baculovirus transfer vector pFASTBAC™DUAL and used for transposition of the genes into the bacmid DNA following transformation of vendor-supplied *E. coli* DH10Bac™ cells. Recombinant bacmids are constructed by transposition of a mini-Tn7 element from the pFASTBAC™DUAL donor plasmid to the mini-*att*Tn7 attachment site on the bacmid. The Tn7 transposition functions are provided *in trans* by a helper plasmid (pMON7124) present in the DH10Bac™ cells. The mini-Tn7 contains an expression cassette consisting of a Gm^R gene, a baculovirus specific promoter, a multiple cloning site and an SV40 poly(A) signal inserted between the left and right arms of Tn7. The recombinant bacmid DNA was isolated from small scale *E. coli* DH10Bac™ cultures and used to transfect *S. frugiperda* cells. This DNA-transfection procedure makes use of a synthetic cationic lipid (CELLFECTIN™ reagent) that interacts spontaneously with DNA to form a lipid-DNA complex with complete entrapment of the DNA. This complex fuses with the cell membrane to facilitate the efficient delivery of its contents to the interior of the cell. Viral stocks harvested from the transfected cells were used to infect *S. frugiperda* cells following amplification of the viruses through two stages of amplification in monolayer cultures, recombinant viruses were plaque purified and subjected to a further three stages of amplification in monolayer cultures. Viral titers of ca. 5.5×10^7 pfu/ml were typically obtained.

The recombinant nature of the resultant baculoviruses was verified using a heminested PCR approach. The primers used were designed according to published sequences (Kimura *et al.*, 1990; Moriishi *at al.*, 1996a) and were previously used in mutagenesis and sequencing procedures (Chapter 2). Amplicons of the expected sizes of 1.4 kb and

1.2 kb were observed for the respective PCR steps. This approach is based on the assumption that any non-specific amplicons produced during the first PCR step should not be able to function as target DNA during the second PCR step due to a lack of complementarity with the inner primer sequence. The internal sense primers (Cmut-F and Dmut-F) used in this study anneal at a position ca. 250 bp from the original sense primers (BNTseqC4 or BNTseqD4) and resulted in the amplification of a 1.2 kb amplicon.

S. frugiperda cells were infected with recombinant and wild-type baculovirus at a MOI of 5 pfu/cell. Fractionated cell extracts were subjected to SDS-PAGE analysis under reducing conditions. Protein samples were first denatured by heating in the presence of a reducing agent and the strong anionic detergent, SDS. The denatured proteins bind SDS and acquire a negative charge. The amount of SDS bound is proportional to the molecular weight of the protein but independent of its amino acid sequence, and thus migration of the SDS-polypeptide complex during electrophoresis is dependent on the size of the polypeptide. However, post-translational modification of a protein, such as glycosylation or phosphorylation may affect its migration, so that the apparent molecular weight is not necessarily a true reflection of the mass of a polypeptide (O'Reilly *et al.*, 1992). During this study, unique protein bands corresponding to the size of the neurotoxins could not be seen in Coomassie brilliant blue-stained gels. These results correspond with those obtained for the expression of botulinum neurotoxin genes in *E. coli* (Clayton *et al.*, 1995; Kiyatkin *et al.*, 1997).

The botulinum neurotoxins are synthesized as 150 kDa single-chain polypeptides that are proteolytically cleaved to generate di-chain neurotoxins composed of a heavy chain (H, 100 kDa) and a light chain (L, 50 kDa) bridged by a disulfide bond. The disulfide bond can be reduced by 2-marcaptoethanol to result in the dissociation of the botulinum neurotoxins into a heavy and light chain (Oguma *et al.*, 1981; Ochanda *et al.*, 1984; Moriishi *et al.*, 1989). The anti-sera used for the immunoblot detection of the expressed BoNT/C_{ST} and BoNT/MutC_{ST} were prepared against the L chain of BoNT/C_{ST} (anti-C_{ST}(L) (Oguma *et al.*, 1980) and full-length BoNT/C_{ST} (anti-C_{ST}) (C. Ellis), respectively. During this study, anti-C_{ST}(L) cross-reacted with the H chain of reduced BoNT/C_{ST} and it is therefore evident that these two chains share common epitopes. The anti-sera used for the detection of the expressed BoNT/D_{SA} and BoNT/MutD_{SA} were prepared against full-

length BoNT/D_{1873} (anti-D_{1873})(Oguma *et al.*, 1981) and full-length BoNT/D_{SA} (anti-D_{SA})(C. Ellis). The anti-D_{1873} has been shown to cross-neutralize the type D_{SA} neurotoxin (Oguma *et al.*, 1981). The type D botulinum neurotoxins display a 96% amino acid homology in their N-terminal regions, 92% amino acid homology in their core regions, and a 30% amino acid homology for their C-terminal regions (Moriishi *et al.*, 1996a).

Western blot analysis of the Baculo-BoNT/C- and Baculo-BoNT/MC- infected *S. frugiperda* cells extracts indicated immunoreactive material migrating at a position that may correspond to the type C toxin light chain. Immunoreactive material migrating at ca. 60 kDa and 69 kDa was detected in fractions of *S. frugiperda* cells infected with Baculo-BoNT/C and Baculo-BoNT/MC, respectively. The difference in migration between the immunoreactive material of BoNT/C_{ST} and BoNT/MutC_{ST} may be due to the amino acid substitutions made in the L chain in order to obtain the mutant forms of the wild-type gene. However, immunoreactive material migrating at a position of ca. 60 kDa is inconsistent with observations for the L chain from native BoNT/C_{ST} migrating at ca. 50 kDa. However, post-translational modifications (e.g. glycosylation or phosphorylation) performed in the eukaryotic environment of the baculovirus expression system has been shown to affect a proteins migration (O'Reilly *et al.*, 1992). These bands were not present in the fractions from the wild-type baculovirus-infected and mock-infected *S. frugiperda* cells.

The Western blot analysis of the Baculo-BoNT/D- and Baculo-BoNT/MD- infected *S. frugiperda* cells were inconclusive due to the high amount of background bands. No specific immunoreactive material was detected for fractions of *S. frugiperda* cells infected with Baculo-BoNT/D or Baculo-BoNT/MD. Pre-adsorbed anti-sera were used in order to eliminate background bands. However, results obtained were similar to those for normal anti-sera. Further serological-based methods (e.g. ELISA and mouse bioassays) will have to be employed in order to prove without doubt the expression of the botulinum neurotoxin genes.

The successful expression of the respective botulinum neurotoxin genes in BAC-to-BAC™ baculovirus system would have allowed for the evaluation of these expressed proteins as recombinant botulinum vaccine candidates. However, the baculovirus expression system does not seem to be a suitable alternative for the expression of

botulinum neurotoxin genes. Other expression systems or strategies (e.g. low G+C content gram-positive organisms) should be investigated in order to achieve higher levels of expression. Some of these options will be discussed in the following concluding chapter.

CHAPTER 4

CONCLUDING REMARKS

The need for immunization agents quickly became apparent once the economic importance of botulism in livestock was realized. The botulinum vaccine currently in use in South Africa rely on problematic anaerobic cultivation of the *Clostridium* bacterium followed by isolation, purification and inactivation of the toxin by treatment with formalin. In contrast, recombinant biotechnology may facilitate the development of a cheaper and more efficacious botulinum vaccine. Over the past decade several advances has been made in the development of new generation recombinant botulinum vaccines. In this study, the full-length genes encoding the type C and D botulinum neurotoxins were amplified, cloned, mutated with PCR-based site-directed mutagenesis, sequenced, and finally expressed in the BAC-TO-BAC^{TM} baculovirus expression system.

The relationship between toxicogenicity and bacteriophage presence in type C and D cultures has long been established (Inoue and Iida, 1968; Eklund *et al.*, 1971) and it has been confirmed that type C and D phages carry the structural genes for the respective neurotoxins (Fujii *et al.*, 1988). During this study, the presence of these bacteriophages in the *C. botulinum* cultures was verified by transmission electron microscopy. Isolated bacteriophage DNA was used as template for the amplification of the full-length genes by PCR amplification using a long-range enzyme developed for long and accurate amplification. The gel-purified PCR amplicons were cloned into the pGEM®-T Easy vector (Promega) and the integrity of the cloned full-length botulinum neurotoxin genes was determined by automated sequencing. The nucleotide sequences of the full-length genes were compared to the published sequences for the respective genes and homologies of 99.8% (BoNT/C_{ST}) and 99.7% (BoNT/D_{SA}) were observed. Deduced amino acid sequences showed an open reading frame without fortuitous termination signals. This is the first report on the direct amplification of the full-length botulinum neurotoxin genes as other approaches have relied on the reconstruction of these genes with PCR fragments generated from culture supernatants and phage DNA libraries (Kimura *et al.*, 1990; Moriishi *et al.*, 1996a,b).

The cloned full-length wild-type neurotoxin genes were used as template DNA in a PCR-based site-directed mutagenesis procedure. Mutations were generated in the zinc-binding motif (His-Glu-aa-aa-His) located in the central region of the light chain sub-unit by an inverse PCR approach. A pair of inverted tail-to-tail primers was designed according to the site of the desired mutation and was subsequently used to amplify the entire double-stranded plasmid containing the cloned gene. An unique *KpnI* restriction site was incorporated into the 5' ends of both the primers to introduce base pair substitutions and thus resulting in mutation of two of the three targeted amino acids. Sequence analysis of the mutant site indicated that the amino acids in the zinc-binding motif (His-Glu-aa-aa-His) were substituted with three different amino acid residues (Gly-Thr-aa-aa-Asn). The rational for this approach was based on a previous report indicating that the substitution of the three amino acids essential for zinc-binding resulted in the complete detoxification of a modified recombinant type C toxin (Kityatkin *et al.*, 1997). This recombinant mutant protein lacked toxicity *in vivo*, failed to block neuromuscular transmission *in vitro* and had no ability to cleave syntaxin in a broken-cell assay. The non-toxic mutant type C toxin retained the ability to translocate from the gut to the general circulation and to evoke protective antibodies (Kityatkin *et al.*, 1997). The development of such a non-toxic botulinum neurotoxin vaccine candidate will eliminate the need for laborious and time-consuming toxoiding procedures. Furthermore, the obtained non-toxic material can be produced without high-containment facilities.

The wild-type and mutant full-length botulinum neurotoxin genes were subsequently cloned into the baculovirus transfer vector pFASTBAC™DUAL and used for transposition of genes into bacmid DNA contained in *E. coli* DH10Bac™ cells. The recombinant bacmid DNA was isolated from *E. coli* DH10Bac™ cultures and transfected into *S. frugiperda* cells. Recombinant viruses were amplified, plaque-purified and subjected to a further three rounds of amplification in monolayer cultures. The presence of the botulinum neurotoxin genes within the bacmid DNA of the recombinant baculoviruses was verified using a semi-nested PCR approach. *S. frugiperda* cells were subsequently infected with the recombinant and wild-type baculoviruses. Cytoplasmic extracts of infected *S. frugiperda* cell lysates were subjected to SDS-PAGE analysis under reducing conditions. Although, no unique proteins corresponding to the size of the botulinum neurotoxins could be observed in Coomassie brilliant blue-stained gels, Western blot analysis, indicated the presence of immunoreactive material for wild-type and mutant type C neurotoxin. This material migrated at a position that may correspond to the type C toxin light chain. The cell extracts prepared from the wild-type baculovirus-infected and mock-infected *S.*

frugiperda cells did not exhibit these bands. The Western blot analysis of the Baculo-BoNT/D- and Baculo-BoNT/MD- infected *S. frugiperda* cells were inconclusive due to the high amount of background bands. Other serological methods (e.g. ELISA or mouse bioassays) should be used to determine if the botulinum neurotoxins were successfully expressed.

The successful high level expression of the type C and D botulinum neurotoxin is necessary for further studies to evaluate the effectiveness of the introduced mutation for the inactivation of the neurotoxins as well as to evaluate the potential of the expressed proteins as recombinant vaccine candidates. The baculovirus expression system does not seem to be an adequate alternative for the expression of botulinum neurotoxin genes. Other expression systems such as *E. coli* and the low G+C content gram-positive organisms (e.g. *Bacillus*) should be investigated in order to obtain higher expression levels.

E. coli expression offers a low cost alternative, despite variable results obtained for the expression of BoNT genes in this system. The expression of botulinum neurotoxin genes in *E. coli* has been a challenge because of codon bias of these extremely AT-rich (ca. 73%) genes (LePenotiere *et al.,* 1995). In order to overcome low expression levels in *E. coli*, a synthetic BoNT/A(Hc) gene fragment was constructed by taking into account the single codon preferences of *E. coli.* Rare codons have been shown to greatly diminish expression levels of recombinant protein in *E. coli* because of translational stalling (Kane, 1995; Kurland and Gallant, 1996). An alternative to increase expression levels is the co-transfection of *E. coli* with a plasmid carrying the gene for the rare codon's cognate tRNA. The "RIG" plasmid contains the genes that encode the three tRNAs (Arg, Ile, Gly) which recognize the codons rarely used by *E. coli* in highly expressed genes. This plasmid has previously been used for the successful high-level expression of AT-rich *Plasmodium* genes (Baca and Hol, 2000). Codon frequency analysis (Nakamura *et al.*, 1998) showed that 8.2% of the codons encoded in the *Plasmodium falciparum* genome are in fact codons which are rarely used in *E. coli.* These codons are the rare aga/agg codons for Arg, ata for Ile, and gga for Gly. Similarly, codon frequency analysis (this study) showed that 10.8% (BoNT/C_{ST}) and 9.3% (BoNT/D_{SA}) of the codons used in these neurotoxin genes are rarely recognised in *E. coli.* The "RIG" plasmid may be applicable for the overexpression of the AT-rich botulinum neurotoxin genes.

E. coli is not considered an ideal host for the expression of vaccine candidates due to toxic cell wall pyrogens. However, these pyrogens should not pose a serious problem if recombinant

botulinum proteins are expressed at high levels. Recombinant botulinum neurotoxin proteins are highly efficacious immunogens (Clayton *et al.*, 1995; Kiyatkin *et al.*, 1997; Smith, 1988) and only a low dose is required to elicit protective immunity against challenges of native toxin as high as 10^6 LD_{50} (Smith, 1998). Furthermore, the simplicity and low cost of the *E. coli* expression system allows for additional purification steps in order to remove toxic cell wall pyrogens, if necessary.

In conclusion, I remain convinced that recombinant biotechnology may provide solutions to problems regarding botulinum vaccine production and facilitate the development of a low cost efficacious botulinum vaccine. However, it is imperative that solutions to low expression levels and neurotoxin instability are sought and found, before recombinant botulinum vaccines will become a reality.

REFERENCES

Baca, A.M., and Hol, W.G.J. 2000. Overcoming codon bias: A method for high-level overexpression of *Plasmodium* and other AT-rich parasite genes in *Escherichia coli. International Journal for Parasitology* 30:113-118

Balbas, P., and Bolivar, F. 1990. Design and construction of expression plasmid vectors in *Escherichia coli. Methods in Enzymology* 185: 14-37

Bengston, I.A. 1922. Preliminary note on a toxin-producing anaerobe isolated from the larvae of *Lucilia caesar. Public Health Report* 37:146-170

Beuken, E., Vink, C., and Bruggeman, C.A. 1998. One-step procedure for screening recombinant plasmid by size. *BioTechniques* 24:748-750

Billman-Jacobe, H. 1996. Expression in bacteria other than *Escherichia coli. Current Opinions in Biotechnology* 7:500-504

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

Birnboim, H.C., and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* 7:1514-1523

Boyle, J.S., and Lew, A.M. 1995. An inexpensive alternative to glassmilk for DNA purification. *Trends in Genetics* 11:8

Burke, G.S. 1919. Notes on *Bacillus botulinus. Journal of Bacteriology* 4:555-565

Chung, C.T., and Miller, R.H. 1988. A rapid and convenient method for the preparation and storage of competent bacterial cells. *Nucleic Acids Research* 16:3580

Clark, J.M. 1988. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Research* 16:9677

Clayton, M.A., Clayton, J.M., Brown, D.R., and Middlebrook, J.L. 1995. Protective vaccination with a recombinant fragment of *Clostridium botulinum* neurotoxin serotype A expressed from a synthetic gene in *Escherichia coli*. *Infection and Immunity* 63:2738-2742

Davies, A.H. 1995. Advances in the use of recombinant baculoviruses for the expression of heterologous proteins. *Current Opinion in Biotechnology* 6:543-547

Dolman, C.E., and Chang, E. 1972. Bacteriophages of *Clostridium botulinum*. *Canadian Journal of Microbiology* 18:67-76

Eklund, M.W., Poysky, F.T., Reed, S.M., and Smith, C.A. 1971. Bacteriophage and the toxigenicity of *Clostridium botulinum* type C. *Science* 172:480-482

Eklund, M.W., and Poysky, F.T. 1974. Interconversion of Type C and D strains of *Clostridium botulinum* by specific bacteriophages. *Applied Microbiology* 27:251-258

Ernst, W.J., Grabherr, R.M., and Katinger, H.W.D. 1994. Direct cloning into the Autographa californica nuclear polyhydrosis virus for generation of recombinant baculoviruses. *Nucleic Acids Research* 22:2855-2856

Fu, F-N., Lomneth, R.B., Cai, S., and Singh, B.R. 1998. Role of Zinc in the structure and toxic activity of Botulinum neurotoxins. *Biochemistry* 37:5267-5278

Fujii, N., Oguma, K., Yokosawa, N., Kimura, K., and Tsuzuki, K. 1988. Characterization of Bacteriophage Nucleic Acids obtained from *Clostridium botulinum* types C and D. *Applied and Environmental Microbiology* 54:69-73

Giménez, D.F., and Ciccarelli, A.S. 1970. Another type of *Clostridium botulinum*. *Zentralblatt für Bakteriologie Mikrobiologie, und Hygiene A* 215:221-224

Gunnison, J.B., and Meyer, K.F. 1929. Cultural study of an international collection of *Clostridium botulinum* and *parabotulinum*. *Journal of Infectious Diseases* 45:119-134

Hall, J.D., McCroskey, L.M., Pincomb, B.J., and Hatheway, C.O. 1985. Isolation of an organism resembling *Clostridium barati* which produces type F toxin from an infant with botulism. *Journal of Clinical Microbiology* 21:654-655

Hanahan, D., Jessee, J., and Bloom, F.R. 1991. Plasmid transformation of *Escherichia coli* and other bacteria. *Methods in Enzymology* 204:63-113

Hauser, D., Eklund, M.W., Kurazono, H., Binz, T., Niemann, H., Gill, D.M., Boquet, P., and Popoff, M.R.. 1990. Nucleotide sequence of *Clostridium botulinum* C1 neurotoxin. *Nucleic Acids Research* 18:4924

Henning, M.W. 1949. *Animal Diseases in South Africa*. Second edition. Central News Agency LTD., South Africa.

Holton, T.A., and Graham, M.W. 1991. A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucleic Acids Research* 19:1156

Imai, Y., Matsushima, Y., Sugimara, T., and Terada, M.B. 1991. A simple and rapid method for generating a deletion by PCR. *Nucleic Acids Research* 19:2785

Inoue, K., and Iida, H. 1968. Bacteriophages of *Clostridium botulinum*. *Journal of Virology* 2:537-540

Inoue, K., and Iida, H. 1970. Inducible bacteriophages in *Clostridium botulinum.* In *Toxic Microorganisms*, edited by Herzberg, M.. U.S. Dept. of Interior, Washington, D.C.

Jones, I., and Morikawa, Y. 1996. Baculovirus vectors for expression in insect cells. *Current Opinion in Biotechnology* 7:512-516

Kane, J.F. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli. Current Opinions in Biotechnology* 6:494-500

Kimura, K., Fujii, N., Tsuzuki, K., Murakami, T., Indoh, T., Yokosawa, N., Takeshi, K., Syuto, B., and Oguma, K. 1990. The complete nucleotide sequence of the gene coding for botulinum type C1 toxin in the C-ST phage genome. *Biochemical and Biophysical Research Communications* 171:1304-1311

Kits, P.A., and Possee, R.D. 1993. A method for producing recombinant baculovirus expression vectors at high frequency. *Biotechniques* 14: 810-814

Kiyatkin, N., Maksymowych, A.B., and Simpson, L.L. 1997. Induction of an immune response by oral administration of recombinant botulinum toxin. *Infection and Immunity* 65:4586-4591

Kriek, N.M., and Odendaal, M.W. 1994. Botulism. In *Infectious Diseases of Livestock with special reference to Southern Africa,* Vol 2, edited by Coetzer, J.A.W., Thomson, G.R., and Tustin, R.C., and associate editor Kriek, N.P.. Oxford University Press, Cape Town, South Africa. pp. 1354 – 1371

Kurland, C., and Gallant, J. 1996. Errors of heterologous protein expression. *Current Opinions in Biotechnology* 7:489-493

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685

LePenotiere, H.F., Clayton, M.A., and Middelbrook, J.L. 1995. Expression of a large, nontoxic fragment of botulinum neurotoxin type A and its use as an immunogen. *Toxicon* 33:1383-1386

Liebman, J.M., LaSala, D., Wang, W., and Steed, P.M. 1999. When less is more: Enhanced baculovirus production of recombinant proteins at very low multiplicities of infection. *BioTechniques* 26:36 – 42

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

Makoff, A.J., Oxer, M.D., Romanos, M.A., Fairweather, N.F., and Ballantine, S. 1989. Expression of tetanus toxin fragment C in *E. coli:* High level expression by removing rare codons. *Nucleic Acids Research* 17:10191-10202

McCroskey, L.M., Hatheway, C.L., Fenicia, L., Pasolini, B., and Aureli, P. 1986. Characterization of an organism that produces type E botulinal toxin but which resembles *Clostridium butyricum* from the feces of an infant with type E botulism. *Journal of Clinical Microbiology* 23:201-202

Minton, N.P. 1995. Molecular Genetics of Clostridial Neurotoxins. In *Clostridial Neurotoxins: The Molecular Pathogenesis of Tetanus and Botulism*, edited by Montecucco, C.. Springer, Berlin, Germany. pp. 161-194

Moller, V., and Scheibel, I. 1960. Preliminary report on the isolation of an apparently new type of *Cl. botulinum*. *Acta Pathologica et Microbiologica Scandinavia* 48:80

Montecucco, C. and Schiavo, G. 1994. Mechanism of action of tetanus and botulinum neurotoxins. *Molecular Microbiology* 13:1-8

Moriishi, K., Syuto, B., Kubo, S., and Oguma, K. 1989. Molecular diversity of neurotoxins from *Clostridium botulinum* type D strains. *Infection and Immunity* 57:2886-2891

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

Moriishi, K., Koura, M., Abe, N., Fujii, N., Fujinaga, Y., Inoue, K., and Oguma, K. 1996(b). Molecular cloning of the gene encoding the mosaic neurotoxin, composed of parts of botulinum neurotoxin types C1 and D, and PCR detection of this gene from *Clostridium botulinum* type C organisms. *Applied and Environmental Microbiology* 62:662-667

Nakamura, Y., Gojobori, T., and Ikemura, T. 1998. Codon usage tabulated from the international DNA sequece databases. *Nucleic Acids Research* 26:334

Ochanda, J.O., Syuto, B., Oguma. K., Iida, H., and Kubo, S. 1984. Comparison of antigenicity of toxins produced by Clostridium botulinum type C and D strains. *Applied and Environmental Microbiology* 47:1319-1322

Oguma, K., Syuto, B., Iida, H., and Kubo, S. 1980. Antigenic similarity of toxins produced by *Clostridium botulinum* type C and D strains. *Infection and Immunity* 30:656-660

Oguma, K., Syuto, B., Agui, T., Iida, H., and Kubo, S. 1981. Homogeneity and heterogeneity of toxins produced by *Clostridium botulinum* type C and D strains. *Infection and Immunity* 34:382-388

Old, R.W., and Primrose, S.B. 1994. Principles of Gene Manipulation: An introduction to genetic engineering. Fifth edition. Blackwell Science, Oxford, UK.

O'Reilly, D.R., Miller, L.K., and Luckow, V.A. 1992. *Baculovirus expression Vectors: a laboratory manual.* W.H. Freeman and Company, New York, USA.

Patel, G., Nasmyth, K., Jones, N. 1992. A new method for the isolation of recombinant baculovirus. *Nucleic Acids Research* 20:97-104

Romanos, M.A., Makoff, A.J., Fairweather, N.F., Beesley, K.M., Slater, D.E., Rayment, F.B., Payne, M.M., Clare, J.J. 1991. Expression of tetanus toxin fragment C in yeast: gene synthesis is required to eliminate fortuitous polyadenylation sites in AT-rich DNA. *Nucleic Acids Research* 19:1461

Romanos, M. 1995. Advances in the use of *Pichia pastoris* for high-level gene expression. *Current Opinion in Biotechnology* 6:527-533

Sambrook, J., Fritsh, E.F., and Maniatis, T. 1989. Molecular cloning– A Laboratory Manual. Second edition. Cold Spring Harbor Laboratory Press, New York.

Seddon, H.R. 1922. The specific identity of *Bacillus parabotulinus*. *Journal of Comparative Pathology and Therapy* 35:275-280.

Simonen, M., and Palva, I. 1993. Protein secretion in *Bacillus* species. *Microbiological Reviews* 57:109-137

Smith, L.A. 1998. Development of recombinant vaccines for botulinum neurotoxin. *Toxicon* 36:1539-1548

Smith, L.DS., and Sugiyama, H. 1988. Botulism: the organism, its toxins, the disease. Second edition. Charles C. Thomas, Illinois, USA.

Sonnabend, O., Sonnabend, W., Heinzle, R., Sigrist, T., Dirnhofer, R., and Krech, U. 1981. Isolation of Clostridium botulinum type G and identification of type G botulinal toxin in humans: reports of five unexpected deaths. *Journal of Infectious Diseases* 143:22-27

Sudbery, P.E. 1996. The expression of recombinant proteins in yeasts. *Current Opinions in Biotechnology* 7:517-524

Suen, J.C., Hatheway, C.L., Steigerwalt, A.G., and Brenner, D.J. 1988. *Clostridium argentinense* sp. nov.: a Genetically homogeneous group composed of all strains of *Clostridium botulinum* toxin type G and some nontoxigenic strains previously identified as *Clostridium subterminale* of *Clostridium hastiforme*. *International Journal of Systematic Bacteriology* 38:375-381

Udaka, S., and Yamagata, H. 1993. High-level secretion of Heterologous Proteins by *Bacillus brevis*. *Methods in Enzymology* 217:23-33

Viljoen, C.C. 1965. Studies of the toxin of *Clostridium botulinum* type C. M.Sc. thesis, Department of Biochemistry, Faculty of Agriculture, University of Pretoria.

Vinet, G., Berthiaume, L., and Fredette, V. 1968. Un bactériophage dans une culture de *Cl. Botulinum*. *Critical Reviews Canadian Biology* 27:73-74