1

1.0

2.8

2.5

3.2

2.2

1.1

2.0

1.8

1.25

1.4

1.6

MicroD

List of Tables

Table I.	<i>Iridoviridae</i> genera	4
Table II.	<i>Escherichia coli</i> strains used.....	20
Table III.	Recombinant plasmid and phage clones used.....	23
Table IV.	Extended restriction site mapping of the TIV genome in the region of the L96 gene: fragments selected by hybridization to 3' and 5' probes derived from L10, and deduced distances from the termini of the TIV insert in L10.....	63
Table V.	Amino acid composition of the L96 gene product	72
Table VI.	Protein products observed in expression studies using the L96 ORF fused to the Bacteriophage T7 RNA Polymerase Promoter.....	104
Table VII.	Possible TIV genomic DNA fragments encoding the L96 gene	106
Table VIII.	Comparison of the FV3 ICR489 gene and the TIV L96 and capsid genes' 5' and 3' UTR regions.....	109

Isolation and Characterization of a Late *Tipula* Iridescent Virus Gene

by

WILLIAM A. HOME, B.Sc. (Hons.)

A thesis submitted to the Faculty of Graduate Studies in
partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

Molecular Genetics Section, Division of Biological Sciences,
NATIONAL RESEARCH COUNCIL OF CANADA
Biology Department, **CARLETON UNIVERSITY**

Ottawa, Canada

April, 1989

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-51199-0

The undersigned hereby recommend to the Faculty of Graduate Studies the acceptance of this thesis, submitted by *William A. Home*, B.Sc. (Hons.), in partial fulfillment of the requirements for the degree of Master Science.

A handwritten signature in black ink, appearing to be 'J.R.A.', followed by a horizontal line.

Chairman, Department of Biology

A handwritten signature in black ink, appearing to be 'M. L. ...', followed by a horizontal line.

Co-supervisor

A handwritten signature in black ink, appearing to be 'S. K. ...', followed by a horizontal line.

Co-supervisor

Carleton University

Abstract

Virions of the cytoplasmic, icosahedral insect virus *Tipula* iridescent virus (TIV) contain approximately 175-225 kbp of dsDNA and 25-30 proteins. Recent molecular genetic studies by Tajbakhsh (1986, 1988) have revealed details on TIV genomic composition and the organization and expression of the major TIV capsid protein. North American TIV contains at least four dsDNA components: L(> 176 kbp) and S1(10.8 kbp) are major components, and S2(6.5 kbp) and S3(~3 kbp) are minor components. The broad objective of this thesis was to provide molecular information on TIV genomic organization. For this purpose the focus was on the characterization of a cloned 11.8 kbp BamHI DNA fragment of L which was previously shown to carry a gene (L96) coding for a 3.6 kb transcript expressed late in TIV-infected *Estigmene acrea* cells (Tajbakhsh, 1988). The 11.8 kbp fragment was extensively mapped with restriction endonucleases, subcloned and DNA sequenced. These data showed that the relevant regions of the 11.8 kbp fragment possess an ORF 2604 nucleotides long encoding a predicted protein of 867 amino acids with an estimated molecular weight of 96 kDa and a pI of 10.91. A hydropathy plot of the putative L96 gene product shows a largely hydrophilic polypeptide with a hydrophobic C-terminus. Evidence for protein production from the L96 gene was obtained by expression in *Escherichia coli* cells and in a coupled *in vitro* transcription and translation system. Comparison of the L96 gene with the TIV capsid gene revealed several putative consensus regulatory sequences.

Acknowledgements

I would like to give special thanks to Dr. Vern Seligy for accepting me into his research group, and for providing continued guidance and support throughout this work. I would also like to thank Drs. Peter Lee and Ken Dimock for advice received over the course of this project.

A special thank you to Mike Dove for his invaluable technical assistance, particularly in the more demanding times. His unselfish willingness to provide help was very much appreciated. I would also like to give special thanks to Dr. Shahragim Tajbakhsh for the gift of clone L10, as well as for enlightening me on many aspects of virology and demonstrating many experimental techniques. Thanks also go out to Drs. Mary Hefford and Ivan Curran for their advice and encouragement. I would also like to thank Dr. W. Sung for synthesizing the oligonucleotides, and Stephen Baird for instructing me in the intricacies of personal computers.

I am especially grateful to the many individuals at the NRC who assisted and encouraged me throughout the course of this work. In particular I would like to thank Terry Dowhanick, Steve Bottega, Peter Foster, Kresh Prakash, Janice Leblanc, Gilles Pilate, and Phil Shwed for their support.

I would like to thank NSERC for providing financial assistance during the course of this study.

Finally, I would like to thank Wendy Simons for her continued belief in me throughout the course of this work.

Table of Contents

Title	i
Acceptance page	ii
Abstract	iii
Acknowledgements.....	iv
Table of Contents	v
List of Tables.....	xi
List of Figures	xii
List of Abbreviations.....	xiv
I. INTRODUCTION	1
A. Discovery and early research on TIV.....	1
B. Classification.....	3
C. Characteristics of TIV.....	5
1. Physical and structural properties	5
a) Chemical components	5
b) Structure	5
c) Genome characteristics and organization	6
d) Protein components.....	7
e) TIV associated enzymes.....	8
2. Post-translational modifications	9
3. Host range	9
D. Pathogenesis: pathway of TIV infection.....	9
1. Symptoms of infection and pathology.....	9
2. Cytopathology and replication	10
3. Virus particle assembly.....	12
E. Molecular information available from related viruses	12
1. Genome organization and replication	12

2. Iridovirus associated enzymes.....	13
3. Transcription and RNA characteristics.....	14
F. Research objective	17
II. MATERIALS AND METHODS.....	19
A. Growth of bacteria.....	19
1. Growth media	19
2. Strains utilized and their propagation	19
B. Description of cloning vectors used.....	21
1. Plasmids	21
2. Bacteriophages.....	21
3. Recombinant clones.....	22
C. Propagation of bacteriophages.....	22
1. M13.....	22
a) Preparation of phage containing supernatant	22
b) Phage growth and plating.....	22
c) Serial dilutions.....	22
2. Lambda gt11.....	24
a) Preparation of phage stock.....	24
b) Phage growth and plating.....	24
c) Serial dilutions.....	24
d) <i>In vitro</i> packaging of lambda gt11 DNA	25
D. Bacterial transformation	25
E. Isolation and characterization of plasmid and phage DNA	26
1. Plasmids	27
a) Purification of DNA fragments via extraction with organic solvents and ethanol precipitation.....	27
b) Extraction of DNA fragments from agarose gels(electroelution)	27

c) Removal of contaminants from DNA by spin columns	28
d) Small scale plasmid preparation.....	28
e) Restriction endonuclease digestion and agarose gel electrophoresis	30
f) 5' and 3' DNA terminal modifications	31
(1) 5' overhangs.....	31
(2) 5' dephosphorylation.....	31
g) DNA ligation	32
h) Radiolabelling of DNA and hybridization.....	32
(1) Nick translation	32
(2) Colony hybridization.....	33
(3) Southern blotting.....	34
2. Bacteriophages.....	35
a) M13.....	35
(1) Extraction and purification.....	35
(a) RF DNA preparation.....	35
(b) Viral "+" strand DNA preparation.....	35
(2) Agarose gel electrophoresis.....	37
(a) Phage DNA.....	37
(b) C-test.....	37
(c) Insert sizing using Nuclease S1	37
(3) Cloning and subcloning	38
(4) Construction of deletion clones for DNA sequencing.....	38
b) Lambda gt11.....	40
(1) Extraction and purification.....	40
(2) Cloning and subcloning	41
(3) Plaque hybridizations	41
F. Isolation and characterization of TIV	41

1. Isolation of TIV virions from infected larvae.....	41
2. Isolation of total TIV DNA from purified virions.....	42
G. Analysis of DNA by nucleotide sequencing.....	43
1. Reactions	44
a) DNA synthesis using DNA-dependent DNA polymerases.....	44
(1) Klenow fragment of <i>E. coli</i> DNA polymerase I	44
(2) T7 DNA polymerase (Sequenase)	45
b) ssDNA sequencing using M13 "+" strand as template	45
c) dsDNA sequencing using plasmid as template.....	45
2. Polyacrylamide gel electrophoresis	46
a) Wedge gels.....	46
b) Uniform thickness gels.....	47
c) Gel autoradiography	47
3. Sequence analysis	48
H. Heterologous expression in <i>E. coli</i>	48
1. Construction of expression clones.....	48
2. Sample preparation and analysis.....	48
a) Protein sample preparation.....	48
(1) pUC 19 based expression clones.....	48
(2) pGEM-2 based expression clones.....	50
(a) Preparation of unlabelled proteins	50
(b) Preparation of ³⁵ S labelled proteins.....	51
b) SDS-PAGE analysis of samples.....	52
c) Western blotting.....	52
I. <i>In vitro</i> synthesis of RNA and protein	54
1. Synthesis of RNA <i>in vitro</i>	54
a) <i>In vitro</i> transcription.....	54

b) Agarose gel analysis of RNA.....	55
2. <i>In vitro</i> translation	55
3. SDS-PAGE analysis of ³⁵ S labelled proteins.....	55
III. RESULTS	57
A. History of the L10 plasmid carrying the L96 gene.....	57
B. Physical mapping of the L96 gene.....	59
1. Restriction site map of clone L10.. ..	59
2. Mapping of the L96 gene on the L component of the TIV genome	59
C. Analysis of the L96 gene.....	65
1. Preparation of deletion clones for DNA sequencing.....	65
2. DNA sequence	68
a) The L96 ORF.....	68
b) Sequence characteristics	68
3. Protein sequence	71
D. Construction of expression clones	71
1. pUC 19 based constructs.....	73
a) Unfused construct	73
b) 5' fusion construct	76
c) "Sandwich" constructs.....	77
2. Lambda gt11 based constructs	82
3. pGEM-2 based constructs.....	84
E. Heterologous expression of the L96 gene in <i>E. coli</i>	88
1. pUC 19 based constructs.....	88
a) Unfused construct	88
b) 5' fusion construct	88
c) "Sandwich" constructs.....	90
2. pGEM-2 based constructs.....	96

a) Unlabelled proteins.....	96
b) <i>In vivo</i> labelled proteins	98
F. <i>In vitro</i> analysis of the L96 gene.....	100
1. <i>In vitro</i> transcription.....	100
2. <i>In vitro</i> translation	100
IV. DISCUSSION	105
A. Restriction site mapping of the TIV genome	105
B. Molecular characterization of the late TIV gene L96	108
1. Sequence characterization of the L96 gene.....	108
2. Expression of L96 in <i>E. coli</i> cells	111
3. <i>In vitro</i> analysis of the L96 gene.....	118
C. Conclusions	119
V. REFERENCES.....	121
APPENDIX A.....	131
APPENDIX B.....	134
APPENDIX C.....	135

List of Tables

Table I.	<i>Iridoviridae</i> genera	4
Table II.	<i>Escherichia coli</i> strains used.....	20
Table III.	Recombinant plasmid and phage clones used.....	23
Table IV.	Extended restriction site mapping of the TIV genome in the region of the L96 gene: fragments selected by hybridization to 3' and 5' probes derived from L10, and deduced distances from the termini of the TIV insert in L10.....	63
Table V.	Amino acid composition of the L96 gene product	72
Table VI.	Protein products observed in expression studies using the L96 ORF fused to the Bacteriophage T7 RNA Polymerase Promoter.....	104
Table VII.	Possible TIV genomic DNA fragments encoding the L96 gene	106
Table VIII.	Comparison of the FV3 ICR489 gene and the TIV L96 and capsid genes' 5' and 3' UTR regions.....	109

List of Figures

Figure 1.	Cloning of the L96 gene from the TIV genome by oligonucleotide hybridization.....	58
Figure 2.	Restriction endonuclease mapping of clone L10.....	60
Figure 3.	Mapping of the L96 gene on the TIV genome.....	62
Figure 4.	Subcloning of L10 restriction endonuclease fragments into M13mp19 for DNA sequencing purposes	66
Figure 5.	Sequencing strategy for the L96 gene.....	67
Figure 6.	Nucleotide sequence and deduced amino acid sequence of the TIV L96 gene and gene product	69
Figure 7.	Location of other open reading frames in the region of the TIV genome near the L96 gene.....	70
Figure 8.	Subcloning of the major portion of the L96 gene.....	74
Figure 9.	Subcloning of the L96 gene into plasmid pUC 19	75
Figure 10.	Fusion of codons 1 to 867 of the L96 ORF to codons 1 to 6 of the alpha peptide of β -galactosidase of pUC 19.....	78
Figure 11.	Regions of the L96 ORF selected for the construction of "sandwich" expression clones.....	80
Figure 12.	Construction of "sandwich" expression clones.....	81
Figure 13.	Fusion of the L96 ORF to the β -galactosidase gene of lambda gt11	83
Figure 14.	Fusing of the L96 ORF to the Bacteriophage T7 RNA Polymerase Promoter	85
Figure 15.	Introduction of a Shine-Dalgarno site 5' of the L96 gene fused to the Bacteriophage T7 RNA polymerase Promoter.....	87
Figure 16.	Expression study of the L96 gene in <i>E. coli</i> cells using constructs with the L96 ORF either fused to its native TIV promoter or to the <i>lac</i> promoter	88

Figure 17.	Expression study of the L96 gene in <i>E. coli</i> cells using "sandwich" constructs	92
Figure 18.	Fractionation of the proteins from the expression studies of the L96 "sandwich" constructs in <i>E. coli</i> into intracellular soluble and insoluble, and extracellular protein fractions.....	93
Figure 19.	Expression study of the L96 gene in <i>E.coli</i> cells using Bacteriophage T7 RNA Polymerase promoter elements	97
Figure 20.	<i>In vivo</i> labelling and expression of the L96 gene in <i>E. coli</i> cells using Bacteriophage T7 RNA Polymerase promoter elements	99
Figure 21.	Analysis of <i>in vitro</i> transcription products made from the L96 gene using Bacteriophage T7 RNA Polymerase.....	101
Figure 22.	Analysis of <i>in vitro</i> translation product produced from the L96 gene.....	103

List of Abbreviations

A	adenine
Å	Angstrom
aa	amino acid
Amp ^r	ampicillin resistant
ASFV	African swine fever virus
ATP	adenosine triphosphate
bp	base pair
C	cytosine
Ci	Curie
CIP	Calf Intestinal Alkaline Phosphatase
CIV	<i>Chilo</i> iridescent virus
cm	centimeter
cm ³	cubic centimeter
cpm	counts per minute
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
dTTP	deoxythymidine triphosphate
dNTP	deoxynucleotide triphosphate
ddATP	dideoxyadenosine triphosphate
DMF	dimethyl formamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
ds	double-stranded

dsDNA	double-stranded DNA
DTT	dithiothreitol
E	early
<i>E. coli</i>	<i>Escherichia coli</i>
EtBr	ethidium bromide
FLDV	fish lymphocystis disease virus
FV3	frog virus 3
g	gram
G	guanine
h	hour(s)
HOAc	acetic acid
ICDV	icosahedral cytoplasmic deoxyribovirus
ICTV	International Committee on Taxonomy of Viruses
IE	immediate-early
IPTG	Isopropylthio- β -D-galactoside
IV	iridescent virus
Kan ^r	kanamycin resistant
kb	kilobase
kbp	kilobase pair
kDa	kilodalton
KOAc	potassium acetate
L	large DNA component of TIV
L	late
<i>lacZ</i>	a portion of the β -galactosidase gene (alpha peptide) located on pUC 19
M	molar
mA	milliampere
MCS	multiple cloning site

mg	milligram
min	minute(s)
MIV	mosquito iridescent virus
ml	milliliter
mm	millimeter
mM	millimolar
mRNA	messenger RNA
N	normal
Na ₂ EDTA	disodium N,N-ethylenediaminetetraacetic acid
NaOAc	sodium acetate
NH ₄ OAc	ammonium acetate
nm	nanometer
nt	nucleotide
NTP	nucleotide triphosphate
°C	degrees Celsius
ORF	open reading frame
³² P	radioactive isotope of phosphorus
PAGE	polyacrylamide gel electrophoresis
PEG	polyethylene glycol
pfu	plaque forming units
p.i.	post infection
pI	isoelectric point
PMSF	phenylmethylsulfonyl fluoride
PTIS	portable translation initiation site
RF	replicating (dsDNA) form of M13 phage
RNA	ribonucleic acid
RNase	ribonuclease

rpm	revolutions per minute
RT	room temperature
³⁵ S	radioactive isotope of sulfur
S	small DNA component of TIV
SDS	sodium dodecyl sulphate
sec	second(s)
SIV	<i>Sericesthis</i> iridescent virus
ss	single-stranded
SSC	saline sodium citrate (1X: 0.15 M NaCl, 0.015 M sodium citrate)
ssDNA	single-stranded DNA
STE	10 mM Tris·Cl pH 8, 1 mM Na ₂ EDTA, 100 mM NaCl
SV40	Simian virus 40
T	thymine
TAE	40 mM Tris-acetate, 2 mM Na ₂ EDTA pH8
TBE	89 mM Tris-borate, 89 mM boric acid, 2 mM Na ₂ EDTA pH 8
TBST	10 mM Tris·Cl pH 8, 150 mM NaCl, 0.05% Tween 20
TC	Top Component of a virus preparation fractionated on a gradient
TE	10 mM Tris·Cl pH 7.6, 1 mM Na ₂ EDTA
TE [*]	10 mM Tris·Cl pH 7.6, 0.1 mM Na ₂ EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
Tet ^r	tetracycline resistant
TIV	<i>Tipula</i> iridescent virus
Tris	tris(hydroxymethyl)aminoethane
U	uracil
U/ul	units per microliter
uCi	microcurie
ug	microgram

um	micrometer
uM	micromolar
v	volt
VC	viroplasmic center
w	watt
w/v	weight/volume

I. INTRODUCTION

A. Discovery and early research on TIV

Viruses constitute a class of infectious microorganisms and are among the smallest of all life forms. As a result of these facts viruses are obligate parasites with a requirement for intracellular growth and a heavy dependence on host cell structural and metabolic components for their replication. This property of viruses has made them invaluable in unravelling the many complexities involved in host cell development. The study of viruses has provided a useful framework for building our understanding of the subtleties in genetics, biochemistry, and cellular and molecular biology. In fact, much of our current understanding of molecular biology and the techniques currently in use in this field were developed during the course of investigating specific viral functions.

In 1954, while conducting studies on the nuclear and cytoplasmic polyhedroses of the European crane fly *Tipula paludosa* (Diptera), Xeros (1954) reported an insect pathogen which caused a cytoplasmic disease of the fat body of *T. paludosa* larvae. Infected larvae were easily detected in strong sunlight as the fat body appeared iridescent blue through the integument. Death was found to occur 2 to 4 weeks after the disease became detectable, and the moribund insect was observed to be purplish-white in colour due to the disruption of the diseased fat cells and the liberation of their contents into the hemolymph of the insect. Using the technique of high speed differential centrifugation on infected fat bodies Xeros was able to obtain a purified viral pellet which was orange in transmitted light and iridescent purple, blue, green, etc. in reflected light. This pellet was composed of approximately 100 nm diameter particles which were responsible for the optical properties of the pellet. Acid hydrolysis of the pellet yielded a product possessing a typical nucleic acid spectrum with a peak at approximately 265 Å. Cytological examination of the infected fat cells showed the iridescent parts of the cell to be intensely Feulgen positive. Together these two facts prompted Xeros to propose that the particles contained DNA. Xeros was also able to

determine that the viral particles were not simple spheres, but were somewhat irregular in shape and composite with many smaller subunit particles being embedded in each virion.

The cytoplasmic site of replication and use of DNA as genetic material of this virus was confirmed by Williams and Smith (1957). These authors were the first to report the astonishingly large quantities of this virus (up to 25% of the dry larval weight) which were produced in infected larvae. The iridescent nature of the virus and infected tissues was explained by Bragg reflections (Williams and Smith, 1957; Klug *et al.*, 1959) which involve the diffraction of light by the viral crystals in a manner analogous to that of X-rays by ordinary crystals. Specifically, Bragg reflections are constructive interference where the wavelength reflected, which is a function of the distance between successive reflecting planes (interparticle distance) in the crystals, is visible. Williams and Smith (1957) were the first to use the term "iridescent" to describe the turquoise coloration of systematically infected larvae and pellets of this virus, and in 1958 named the virus *Tipula* iridescent virus (TIV).

Early structural studies (Williams and Smith, 1957, 1958; Smith, 1958; Klug *et al.*, 1959) established that TIV possessed a regular icosahedral shape. Due to the ease of purification of large quantities of TIV (Williams and Smith, 1957) pioneering studies which lead to much of our basic understanding of the structure of icosahedral viruses were conducted on TIV (Caspar and Klug, 1962).

TIV was the first insect virus observed to be transmitted not only from one species to another within an order, but between three different orders of insects: Diptera (true flies); Lepidoptera (butterflies and moths); and Coleoptera (beetles) (Smith and Rivers, 1959; Bird, 1961; Smith *et al.*, 1961). In all cases the larval fat body was the initial site of virus development; however, the virus also multiplied in skin, muscles, wing buds, legs, and the head (Smith *et al.*, 1961). In all tissues the cellular site of viral replication was the cytoplasm (Smith *et al.*, 1961). Different species of insects were observed to yield different colored iridescence in the viral pellet--presumably due to either slight variations in particle size or in

the packaging of the microcrystals. The very wide host range of TIV is another feature which prompted further studies on this interesting virus.

TIV was shown by Thomas and Williams (1961) to consist of an outer protein shell surrounding an inner nucleoprotein core. The inner core composition was estimated at 26-34% DNA and 60-65% protein.

B. Classification

In 1976 the International Committee on Taxonomy of Viruses (ICTV) created the family *Iridoviridae* to encompass several different vertebrate and invertebrate viruses which did not fit into any of the other established groups. The unifying features of this new virus family were: 1) polyhedral symmetry; 2) large (approximately 170 kilobase pairs), linear, double stranded (ds) DNA genome; and, 3) a cytoplasmic site of replication (Matthews, 1982; Willis, 1985). Subsequent research yielded sufficient biochemical data that at the 1982 ICTV meeting the five genera of the family *Iridoviridae* shown in Table I were established (Matthews, 1982).

The name "iridovirus" was derived from the observation that larvae infected with many of the insect viruses, as well as purified pellets of these viruses, glowed with a blue or green iridescence. Presumably this iridescence was due to the aforementioned Bragg reflection of the viral crystals. However, none of the vertebrate "iridoviruses" display this particular characteristic, and for this reason the more descriptive name of icosahedral cytoplasmic deoxyriboviruses (ICDVs) has been proposed (Willis, 1985).

Insect viruses form the largest group within the *Iridoviridae* encompassing the *Iridovirus* (small iridescent virus) and *Chloriridovirus* (large iridescent virus) genera. To date 32 insect iridescent viruses have been isolated (Kelly, 1985). One nomenclature scheme which has been proposed for iridescent viruses is their numbering in order of their chronological discovery (Kelly, 1985). Under this system TIV is referred to as Iridescent Virus Type I (IV1). The best studied member of the *Iridoviridae* family is frog virus 3 (FV3; Willis *et al.*, 1985). Other prominent members include the lone mammalian virus of the family, the

TABLE I. *Iridoviridae* genera¹.

English vernacular name	International name	Type species
1. Small iridescent virus	<i>Iridovirus</i>	<i>Tipula</i> iridescent virus (Type 1)
2. Large iridescent virus	<i>Chloriridovirus</i>	Mosquito iridescent virus (Type 3)
3. Amphibian icosahedral deoxyribovirus	<i>Ranavirus</i>	Frog virus 3 (FV3)
4. Fish lymphocystis disease virus	<i>Lymphocystivirus</i>	Flounder lymphocystis disease virus (FLCDV)
5. African swine fever virus	-	African swine fever virus, Kenya strain (ASFV)

¹ Adapted from Willis, 1985.

African swine fever virus (ASFV) which causes an important disease of domestic pigs (Vinuela, 1985), and fish lymphocystis disease virus (FLDV) which is the causative agent in a chronic disease of several higher order salt and fresh water fish (Flugel, 1985). Recently a large (approximately 190 nm) ICDV that infects certain eukaryotic *Chlorella*-like green algae and has cytopathic properties similar to the other ICDVs has been reported (Van Etten *et al.*, 1983; Meints *et al.*, 1986).

C. Characteristics of TIV

1. Physical and structural properties

a) Chemical components

TIV virions contain 80% protein (Thomas, 1961) distributed between the outer shell and inner core, 12.4-19% DNA (Glitz *et al.*, 1968) found in the core, and 5.2% lipid (Kalmakoff and Tremaine, 1968). No RNA or polysaccharides were found to be associated with the viral particle (Thomas and Williams, 1961).

b) Structure

Members of the *Iridoviridae* family share common structural features. Through the use of techniques such as negative staining and thin sectioning, freeze-etching, and neutron scattering it has been shown that the *Iridoviridae* shell is composed of three concentric domains: a proteinaceous icosahedral capsid, an intermediate unit membrane composed of lipids and proteins, and a central core containing the viral DNA and its associated proteins (Stoltz, 1973; Cuillel *et al.*, 1979; Darcy-Tripier *et al.*, 1984). Klump and coworkers (1983) have demonstrated that the viral DNA within the core is attached to the coat protein to form a chromatin-like structure. TIV virions have a dehydrated diameter of 130 nm (Williams and Smith, 1958), and a hydrated diameter of 250 nm (Klug *et al.*, 1959). The inner core of thin sections of osmium fixed TIV pellets is observed to be 90 nm in diameter (Thomas and Williams, 1961). The particle weight and volume of the TIV virion have been estimated at 2×10^{-15} g and 2.6×10^{-15} cm³, respectively (Thomas, 1961; Deblois *et al.*, 1978).

The outer protein shell of TIV virions has been described as a skewed icosahedron (Wrigley, 1970; Stolz, 1973). The icosahedral structure is composed of 20 equilateral triangular faces consisting of smaller morphological subunits called capsomeres. Manyakov (1977) studied the structure of the TIV icosahedron in detail using negative staining techniques. He observed that the capsid was broken down by long term storage at 4°C or treatment with chloroform and yielded pentagons and triangles as subunits. He proposed a model consisting 12 pentagons, each composed of 31 subunits, and 20 triangles, each consisting of 20 subunits, for a total of 1472 subunits in the TIV icosahedron. This value was identical to that proposed by Wrigley in 1970 after revising his initial prediction of 1562 subunits (Wrigley, 1969).

c) Genome characteristics and organization

The genome of TIV has been estimated at 190-235 kbp from reassociation kinetics (Bellet and Inman, 1967) and 176-247 kbp from restriction endonuclease analysis (Tajbakhsh, *et al.*, 1986). Early reports revealed that TIV DNA possessed a high molar dissymmetry ratio ($A + T/G + C$) of 2.20, and that the DNA was not methylated (Thomas, 1961; Kawase and Hukuhara, 1967; Bellet and Inman, 1967). Other insect iridoviruses such as *Chilo* iridescent virus (CIV) and *Sericesthis* iridescent virus (SIV) were also observed to be A-T rich (Bellet and Inman, 1967). Recent studies involving isolation, cloning, and DNA sequencing of the TIV genome have confirmed these earlier observations (Tajbakhsh *et al.*, 1986; Tajbakhsh, 1988; this thesis). These results differ from those observed with FV3 and FLDV which, through the use of various isoschizomeric restriction enzymes which differentiate between methylated and non-methylated sequences, have been found to be methylated at cytosine residues (Willis and Granoff, 1980; Flugel, 1985).

Delius *et al.* (1984) showed the CIV DNA molecule to be linear, circularly permuted, and terminally redundant. Two other *Iridoviridae* genomes, FV3 and FLDV, have also been shown to be circularly permuted and terminally redundant (Goorha and Murti, 1982; Darai *et al.*, 1983). Teminal redundancy describes the repetition of the same sequence at both ends of

a viral genome (thereby permitting circularization of the genome), while circular permutation is the phenomena of this repeated region not necessarily being identical in any two DNA molecules (i.e. abcdefghijka represents a terminally redundant genome, while bcdefghijkab, cdefghijkabc, defghijkabcd, etc. represent a population of circularly permuted and terminally redundant genomes). Both of these are common features in some viruses such as the bacteriophage T4 (Miller, 1975; Skalka, 1977), and have not been demonstrated for any other eukaryotic viruses.

In contrast to all other iridoviruses which possess monopartite genomes, TIV has been shown to possess a genome consisting of four components (Tajbakhsh *et al.*, 1986). The largest of these components, L (> 176 kbp) and S1 (10.8 kbp), are easily visible in EtBr stained agarose gels; while the smaller fragments, S2 (6 kbp) and S3 (3.5 kbp), are detected by extracting S DNA, re-electrophoresing, and probing a replica blot of the gel with labelled S1 DNA. All three S components are related to L with S1 and S2 being the most similar and S3, which contains some unique sequences, being a subset of S2. Restriction enzyme analysis of TIV DNA also revealed the presence of sub-molar bands (Tajbakhsh *et al.*, 1986). Similar results have been observed with the *Choristoneura fumiferana* multiple nucleocapsid Nuclear Polyhedrosis Virus (MNPV) and were attributed to the presence of variants arising from the passage of plaque-purified virus in larvae (Arif and Doerfler, 1984). Tajbakhsh *et al.* (1986) also reported what appeared to be repetitive sequences in the TIV genome, but these observations could not easily explain the submolar bands observed by restriction endonuclease analysis.

d) Protein components

TIV has been shown to consist of 25-30 polypeptides (Krell and Lee, 1974; Tajbakhsh and Seligy, 1989) ranging in molecular weight from 17,500 to 300,000 (Krell and Lee, 1974). The smallest of these proteins appears to be a lipoprotein, and none appear to be glycoproteins when subjected to the PAS reaction. This latter result agrees with the earlier observation of Thomas (1961) that no neutral or amino polysaccharides were detectable in TIV. The viral

proteins are detectable in PAGE gels of homogenates of TIV infected hemocytes of *Galleria mellonella* larvae obtained 4 hours after viral injection (Krell and Lee, 1974). Six non-capsid viral polypeptides are also observed in infected cells (Krell and Lee, 1974). To date the functions of these polypeptides are largely unassigned, although the 46 kDa polypeptide has been shown to be the major protein involved in the formation of the TIV capsid structure (Tajbakhsh, 1988). The gene encoding this protein has been cloned and DNA sequenced (Tajbakhsh, 1988).

When a crude preparation of TIV virions is centrifuged on a linear sucrose gradient several bands are usually observed (Krell and Lee, 1974). The band of highest density represents intact virions corresponding to infectious virus. A less dense band dubbed "top component" (TC) by Krell and Lee (1974) is located immediately above the main band. A band consisting of empty capsids is found nearest to the meniscus of the gradient. More recent studies using the electron microscope to detect the permeability to negative stains of the viral core have defined these bands as corresponding to "full", "partially filled", and "empty" virions, respectively (Lee and Brownrigg, 1982). Recently Tajbakhsh *et al.* (1986) showed that the TC fraction contained less DNA than full virions, and this DNA contained the highest molar ratio of the S1 to the L DNA component. Krell and Lee (1974) purified TC, subjected it to PAGE analysis, and were able to resolve 12 protein bands. Of these 12 bands 9 were of equal intensities and three were of lesser intensities as compared to those in gels of mature TIV particles.

e) TIV associated enzymes

Certain enzyme activities have been associated with *Iridoviridae* virions including protein kinase, nucleotide phosphohydrolase, and RNA polymerase (Kelly, 1985). However, recent studies (Franke and Hruby, 1986) have shown that vaccinia virus will fortuitously package chloramphenicol acetyl transferase, a bacterial enzyme produced in the eukaryotic cell by recombinant means, into virions by way of "hitchhiking". Tajbakhsh and Seligy (1989) have

suggested that the significance of enzyme activities reported to be associated with TIV virus particles should be questioned as these proteins are not necessarily virally encoded.

2. Post-translational modifications

Studies by Krell and Lee (1974) confirmed the earlier observation of Thomas and Williams (1961) that no TIV proteins are glycosylated. Glycosylation has also not been detected in mosquito iridescent virus (MIV; Wagner *et al.*, 1974) and FV3 proteins (Goorha and Granoff, 1979). Some FV3 core proteins have been reported to be phosphorylated (Aubertin *et al.*, 1980).

3. Host range

The small iridescent viruses frequently have a broad host range when injected; however, a narrower range is observed when entry occurs by a more natural route such as oral ingestion (Hall, 1985). As mentioned previously, TIV not only infects other species within the same order (Diptera, 7 species), but also infects species in two other orders--Lepidoptera (11 species) and Coleoptera (3 species) (Smith and Rivers, 1959; Smith *et al.*, 1961; Bird, 1961). TIV has been shown to be infective to all four larval instars, pupae, and adults of both sexes of *Tipula oleracea*, but not to the embryo (Carter, 1974).

D. Pathogenesis: pathway of TIV infection

1. Symptoms of infection and pathology

The pathogenesis of iridescent viruses is poorly understood, primarily because little is known about how the virus gains access to the initial replication site. The viruses are not highly infectious when administered orally, and probably enter the larvae through wounds, cannibalism, or by parasites (Carter, 1973; Kelly, 1985).

The initial site of TIV multiplication is not known, although the hemocytes are popular cells to study when evaluating the early stages of pathogenesis, and these cells may well be the primary site of replication (Kelly, 1985). The virus then rapidly spreads through the larva. Initially, the fat body is infected (Xeros, 1954; Williams and Smith, 1957; Smith, 1958; Smith *et al.*, 1961), then the virus produces a systemic infection by rapidly multiplying in skin,

muscles, wingbuds, legs, and the head (Smith *et al.*, 1961). The resulting infection renders the larvae flaccid within 7-10 days of infection, although death may take longer (Kelly, 1985). The rate of pathogenesis is temperature dependent, and at optimum temperatures the virus produces the aforementioned systemic infection (Kelly, 1985).

TIV infection causes cell fusion in tissue culture (Mathieson and Lee, 1981), profound alterations in nuclear morphology (Mathieson and Lee, 1981; Kelly, 1985), and rapid inhibition of host cell macromolecular synthesis (Mathieson and Lee, 1981; Lee and Brownrigg, 1982; Kelly, 1985). Interestingly, Lee and Brownrigg (1982) showed that UV-inactivated virus was as efficient as live virus in inhibiting host macromolecular synthesis, but that β -propiolactone treated virus had no such effect. These authors concluded that a virus specific protein either had a direct effect on the nucleus, or acts as a catalyst in the cell for the synthesis of an inhibitor which blocks host macromolecular synthesis.

2. Cytopathology and replication

The bulk of our understanding of insect iridescent virus cell cytopathology comes from infection studies *in vitro*. The general features of replication are common to all of the iridescent viruses, although some variations in the timing of events has been noted (Hall, 1985). The primary point of viral replication and assembly is the cytoplasm (Younghusband and Lee, 1969; Yule and Lee, 1973; Mathieson and Lee, 1981; Kelly, 1985). A nuclear involvement has been implicated in FV3 replication (Goorga *et al.*, 1977, 1978), but not for any other iridoviruses although ASFV fails to replicate in enucleated VERO cells (Ortin and Vinuela, 1977).

Since many insect cells in culture are susceptible to small iridescent virus infection the cell pathology has been mainly studied in hemocytes or cells in culture (Kelly, 1985). As described by Kelly (1985) iridescent virus replication involves entrance by pinocytosis. Subsequently the virus particles have been observed to be taken up into lysosomes (Younghusband and Lee, 1969; Mathieson and Lee, 1981), although it has been claimed that naked virions can be seen in the cytoplasm (Kelly and Tinsley, 1974). Mathieson and Lee

(1981) observed the disintegration of virus particles in vesicles, and this has been interpreted as uncoating.

The next obvious step in the morphogenesis of the virus is the appearance of the viroplasmic centers (VCs) in the cytoplasm of the infected cell (Kelly, 1985). Viroplasmic centers appear as membrane-free isolated pockets of cell cytoplasm containing ribosomes ("loose matrix") engulfed by a homogeneous "dense matrix" resembling chromatin in appearance and devoid of ribosomes (Mathieson and Lee, 1981). Electron microscope autoradiography has localized viral DNA exclusively in the "dense matrix" zones, whereas progeny virions are assembled within the "loose matrix" zones of the VC (Yule and Lee, 1973; Mathieson and Lee, 1981). Dense and loose matrix regions of the VC are also observed in SIV and MIV infections (Mathieson and Lee, 1981). VCs are surrounded by mature virus particles and various forms of incomplete viruses which may be in the course of assembly (Yule and Lee, 1973).

Release of virus from infected cells may occur by exocytosis at the plasma membrane (Mathieson and Lee, 1981), although this appears not to be as extensive a phenomenon as observed with vertebrate iridoviruses (Kelly, 1985). Most likely cell lysis and dispersion of the virus into the hemolymph or cell culture fluid is the most important mechanism of release (Mathieson and Lee, 1981; Kelly, 1985).

Recently the nature of the membrane-less TiV VCs have been investigated (Seagull *et al.*, 1985; Bladon *et al.*, 1986; Bertin *et al.*, 1987). Seagull *et al.* (1985) observed no cytoskeletal changes during the initial stages of TiV infection and concluded that cytoskeletal elements were not important in the formation or maintenance of these membrane-free cytoplasmic virus assembly sites. Subsequently, Bladon *et al.* (1986) were able to isolate intact VCs with a monoclonal antibody against lymphocyte nuclear matrix proteins suggesting the possible involvement of highly conserved nuclear proteins in the assembly and maturation of virions as well as in the maintenance of the integrity of the VCs. Associations with the nuclear matrix have been shown for nuclear DNA viruses such as herpes simplex type I or SV40

(Bibor-Hardy *et al.*, 1982; Ben-Ze'ev and Aloni, 1983; Ben-Ze'ev *et al.*, 1983; Milavek *et al.*, 1984). The lack of microtubule and microfilament involvement in TIV VC formation was conclusively proved by Bertin *et al.* (1987) who showed that taxol, colchicine, or cytochalasin B treatment had no effect on the development of the VCs. These results with TIV differ distinctly from those with FV3 in which cytomatrix filaments, intermediate filaments, and the microtrabecular lattice have all been implicated in the formation and maintenance of the FV3 VCs (Murti and Goorha, 1983; Murti *et al.*, 1984; Murti *et al.*, 1988).

3. Virus particle assembly

Several models have been proposed for the assembly of iridescent viruses. Initial suggestions (Smith, 1956; Hills and Smith, 1959) centered around the formation of viral membranes or shells before development of the core. Bird (1961) proposed development of the core first followed by encapsidation with protein; while Xeros (1964) proposed a parallel development of the viral shell and DNA fibrils. More recently Yule and Lee (1973) put forth a hypothesis of viral assembly combining aspects of particle development suggested by Smith (1956) and Xeros (1964) based on cytological and immunological evidence. This hypothesis involves a sequential construction of the viral shell from the viroplasmic matrix, followed by introduction of the core material into the completed shell, possibly through an opening left in the shell. The use of the cloned capsid gene, as well as other cloned TIV genes, should be able to shed light on the mechanism of TIV particle assembly.

E. Molecular information available from related viruses

1. Genome organization and replication

As previously mentioned the genomes of many iridoviruses (CIV, SIV, TIV, FLDV, ASFV) are A-T rich (Bellet and Inman, 1967; Flugel, 1985; Vinuela, 1985); however, the genome of FV3 has been found to be slightly G-C rich (Willis *et al.*, 1985). Additional variations in *Iridoviridae* genome characteristics occur in the degree of methylation and overall structure. Both FV3 and FLDV DNAs are heavily methylated at cytosine residues (Willis *et al.*, 1985; Flugel, 1985), while others such as TIV and CIV do not appear to be

methyated (Tajbakhsh *et al.*, 1986; Devauchelle *et al.*, 1985). FV3, FLDV, and CIV have all been shown to possess circularly permuted and terminally redundant genomes (Goorha and Murti, 1982; Darai *et al.*, 1983; Delius *et al.*, 1984). In contrast the ASFV genome has been shown to be similar to that of poxviruses (Moss, 1985) in that it contains both terminal and internal inverted repeats and terminal cross-links (Vinuela, 1985). Similar features have not yet been detected in the TIV genome, although work is currently being conducted (Tajbakhsh and Seligy, 1989) to determine if the TIV genome is also circularly permuted and terminally redundant.

Although cytological data are available on iridescent virus replication, including FV3, FLDV, and ASFV, in general molecular information on the replication of these viruses is lacking. Most data comes from work with FV3 (reviewed by Willis *et al.*, 1985) in which it has been shown that DNA replication occurs in both the nucleus and the cytoplasm (Willis *et al.*, 1985). Early in FV3 infection genome to twice genome size daughter molecules are synthesized in the host cell nucleus, while at later times concatemeric DNA is synthesized in the cytoplasm where virus assembly also takes place (Goorha, 1982). Results from pulse-chase experiments showed that this concatemeric DNA served as the precursor for the production of mature FV3 DNA. Also, this concatemeric arrangement is consistent with the finding that FV3 DNA is circularly permuted and terminally redundant, and provides additional evidence for a specific pathway of virus assembly, namely the packaging of DNA into preformed heads (Goorha and Murti, 1982). As mentioned previously ASFV also exhibits a nuclear requirement as it fails to replicate in enucleated VERO cells (Ortin and Vinuela, 1977), although this observation has not yet been explained at the molecular level. To date a nuclear involvement during TIV replication has not been illustrated.

2. Iridovirus associated enzymes

Numerous enzyme functions are observed to be associated with *Iridoviridae* virions. However, as described previously, care must be taken in interpreting these observations. Protein kinase and nucleoside triphosphate phosphohydrolase activities are associated with

FV3, FLDV, and ASFV virions (Willis *et al.*, 1985; Flugel, 1985; Vinuela, 1985). Both FV3 and FLDV have RNase and DNase activities associated with their virus particles (Willis *et al.*, 1985; Flugel, 1985). Somewhat surprisingly, since they both possess methylated cytosine residues in their DNAs, only FV3 but not FLDV has been shown to encode DNA-methyltransferase (Willis *et al.*, 1984b; Flugel, 1985). In the case of FV3 Essani *et al.* (1987) suggest that the DNA-methyltransferase activity is a component of a complicated multifunctional enzyme system that catalyzes diverse functions under appropriate conditions, and is analogous to the type I restriction-modification system in prokaryotes (Yuan, 1981; Bickle, 1982). FV3 virions have also been shown to contain a protein phosphatase (Willis *et al.*, 1985) and a DNA polymerase (Chinchar and Granoff, 1984). In addition to the aforementioned enzymes ASFV is also observed to contain a topoisomerase, a DNA-dependent RNA polymerase, and all the enzymes required for the 5' and 3' end modifications of mRNA (Vinuela, 1985). These latter two groups of enzymes are very similar to those found in poxviruses (Moss, 1985), and lend support to the argument that ASFV does not belong in the *Iridoviridae* family (Vinuela, 1985). To date FLDV is the only iridovirus which has been shown to possess thymidine kinase activity (Flugel, 1985).

3. Transcription and RNA characteristics

Studies with both the cytoplasmic (as exemplified by vaccinia virus, FV3, and ASFV) and the nuclear (as exemplified by herpes virus and the insect nuclear polyhedrosis viruses) ds DNA viruses have revealed a temporal regulation of viral mRNA and protein synthesis in a co-ordinately regulated and sequentially ordered fashion which can be divided into three general classes: immediate-early (IE), early (E), and late (L) depending on whether synthesis occurs before or after DNA replication (Willis *et al.*, 1977; Wood, 1980; Moss, 1985; Roizman and Batterson, 1985; Escribano and Tabares, 1987). This division into three classes is based on differences in the order of gene expression and in the viral gene products that are required for their expression.

The immediate-early class of viral mRNA is distinguishable from subsequent classes via its ability to be synthesized in the absence of *de novo* protein synthesis (i.e. in the presence of protein synthesis inhibitors such as cycloheximide) (Escribano and Tabares, 1987). Early and late genes are dependent on *de novo* protein synthesis for their transcription, and are separated on the basis of their requirement for amplification of the viral DNA, with early gene transcription occurring in the absence of DNA replication and late gene transcription requiring it (Escribano and Tabares, 1987). Early genes generally code for many enzymes, some involved in DNA replication, while late genes code predominantly for structural proteins (Moss, 1978; Horwitz, 1985; Roizman and Patterson, 1985).

Much of the information on the fine structure and regulatory sequences of the viral genes of cytoplasmic viruses has come from studies on vaccinia virus and FV3. Vaccinia virus is a member of the Poxvirus family and contains a ds DNA genome 187 kbp in length. It replicates in the cytoplasm of infected cells and produces mature mRNAs which have 5' caps and 3' poly-A tails. This virus encodes many of the enzymes required for DNA replication and the transcription of early and late genes (Moss, 1985). ASFV has been found to be extremely similar to vaccinia virus in this regard as it contains the enzymes required for mRNA synthesis and produces mature mRNAs which are capped and polyadenylated independently of any host RNA polymerase activity (Vinuela, 1985).

No temporal clustering of vaccinia virus genes has been observed (Venkatesan *et al.*, 1982; Moss, 1985), and, as might be expected, the temporal regulation of vaccinia virus genes has been attributed to the early and late genes possessing different promoters (Weir and Moss, 1987a, 1987b). Since host RNA polymerase II is not used for vaccinia virus transcription, the vaccinia virus promoters and termination sequences are different from those found in eukaryotes (Venkatesan *et al.*, 1981, 1982; Weir and Moss, 1983; Moss, 1985; Rosel *et al.*, 1986; Weir and Moss, 1987a, 1987b; Yuen and Moss, 1987; Miner *et al.*, 1988).

Initial FV3 RNA synthesis occurs in the nucleus using host cell RNA polymerase II (Goorha, 1981), while late viral mRNA synthesis appears to be a cytoplasmic function

possibly under control of a virally encoded enzyme (Willis *et al.*, 1984a). Although it uses host cell RNA polymerase II for at least some of its transcription, FV3 has been found to possess promoters and termination sequences which differ from the consensus sequences of both eukaryotes and vaccinia virus (Willis *et al.*, 1984a; Beckman *et al.*, 1988; Mesnard *et al.*, 1988). Moreover, the two FV3 IE genes cloned and sequenced to date (ICR169 and ICR489) were observed to possess different promoter sequences and lack a classical TATA box (Willis *et al.*, 1984a; Willis, 1987; Beckman *et al.*, 1988; Mesnard *et al.*, 1988). It has been proposed that one or more viral structural proteins interact with the host cell RNA polymerase II and that this interaction somehow permits transcription of the FV3 IE genes by the modified RNA polymerase II (Willis *et al.*, 1984a). Analysis of the only TIV gene sequenced to date (Tajbakhsh, 1988) revealed a CCAAT sequence 129 nt upstream from the ATG, and three A-T rich sequences that could act as putative promoter sites at 39, 55, and 81 nt upstream from the ATG. In comparison only one (ICR489) of the two FV3 genes which have been sequenced was found to contain TATA, CCAAT, and G-C rich sequences similar to eukaryotic motifs (Beckman *et al.*, 1988). These FV3 motifs are similar in appearance and location when compared with the A-T rich and CCAAT sequences of the TIV capsid gene (Tajbakhsh, 1988).

FV3 mRNAs possess the common cap structure at the 5' end, but lack a poly-A tail at the 3' end (Willis and Granoff, 1976; Martin *et al.*, 1982). In contrast, ASFV mRNAs are distinctly different from those of FV3 and are similar to those of vaccinia virus in that they are polyadenylated and are synthesized in the cytoplasm independently of any host cell RNA polymerase II activity (Vinuela, 1985). The work performed on TIV so far has not yet resolved the question of polyadenylation of TIV transcripts, although the TIV capsid gene has been shown to possess a eukaryotic polyadenylation signal at its 3' end (Tajbakhsh, 1988; Tajbakhsh and Seligy, 1989). The polyadenylation of TIV transcripts is currently being studied by comparison of the DNA sequences of genomic and cDNA clones (Tajbakhsh and Seligy, 1989).

F. Research objective

The preceding introduction indicates that a considerable amount of information on the gross structural properties of TIV has been gathered in the three decades since its discovery by Xeros in 1954. These studies have increased our understanding of viral architecture, and have helped in the creation of models to explain the icosahedral capsid structure (Caspar and Klug, 1962). Unfortunately, until recently little or no molecular information has been produced to explain the mode of replication, pathogenesis, and the regulation of gene expression of this and other insect iridescent viruses. The work of Tajbakhsh *et al.* (1986) has shown that the TIV genome possesses some novel features that have not been described for other *Iridoviridae* members, or for other large ds DNA viruses. These features, along with the ease with which large quantities of this virus can be purified from infected larvae, make TIV an excellent candidate to answer some of the questions raised in this section. Hence, the main objective of this thesis was to provide molecular information on genes coded for by this virus.

The information obtained from the isolation and characterization of TIV genes can be utilized to help define the genomic organization of TIV and to evaluate the regulatory mechanisms involved in TIV gene expression in insect cells. Recently, Tajbakhsh (1988) cloned and sequenced the gene encoding the major TIV capsid protein. In the course of this work the approximate locations of several other genes (transcription units) were revealed, including a putative late TIV gene (LW1) located on an 11.8 kbp BamHI TIV DNA fragment. Fine mapping and sequencing of this gene would permit comparison of its flanking sequences with those of the capsid gene to reveal the respective location of these genes in the TIV genome; and, more specifically, to determine the presence of any consensus regulatory sequences in TIV late genes. Moreover, since late viral genes have generally been less accessible to analysis than immediate-early or early genes (Wittek, 1982; Lee-Chen and Niles, 1988), the molecular characterization of a second late TIV gene would provide further information on the temporal regulation of TIV genes. In a broader context, the mapping of a

second TIV gene and its flanking sequence should contribute to our understanding of the gene organization, regulation, and evolution of these large ds DNA viruses. With these considerations in mind, my thesis project was to characterize the 11.8 kbp TIV DNA fragment encoding the LW1 gene, and to detail the molecular features of this putative late gene which is now called L96.

II. MATERIALS AND METHODS

A. Growth of bacteria

1. Growth media

The various liquid media were prepared as described in Appendix A. Agar containing petri dishes of these media were prepared by adding 15 g Bactoagar (Difco) per liter of liquid media prior to sterilization. The resultant hot liquid was cooled to 50°C, and approximately 15 ml aliquots were poured into each sterile petri dish. Minimal media plates were prepared by autoclaving 10.5 g K_2HPO_4 , 4.5 g KH_2PO_4 , 1.0 g $(NH_4)_2SO_4$, and 0.5 g sodium citrate·2H₂O in 100 ml of water. This solution was mixed with a previously autoclaved solution containing 900 ml water and 15 g Bactoagar. Finally, 1.0 ml of filter sterilized 20g/100ml $MgSO_4$, 0.5 ml of filter sterilized 1% thiamine hydrochloride (Sigma), 10 ml of filter sterilized 4 mg/ml histidine (sigma), 10 ml of filter sterilized 4 mg/ml proline (Sigma), and 10 ml of filter sterilized 20% glucose (BDH) were added to this solution and the plates poured.

Antibiotic containing petri dishes were prepared as described for normal petri dishes, except that just prior to pouring the plates an appropriate volume of sterile 100X antibiotic stock solution (Appendix A) was added to the molten medium.

2. Strains utilized and their propagation

The *Escherichia coli* strains used in this work are listed in Table II.

For short term storage HB101 cells were kept on L plates at 4°C, Y1088 and Y1090 cells were kept on L plates containing 30 ug/ml ampicillin at 4°C, while both JM101 and JM109 were kept on minimal media plates at 4°C. In all cases the bacteria were restreaked on fresh plates on a monthly basis. Only HB101 cells were stored on a long term basis in the form of frozen samples. These cells were prepared by diluting an overnight culture of HB101 grown in L medium with one volume of sterile 40% glycerol-60% L medium. Aliquots of 1 ml were flash-frozen by immersion in liquid nitrogen and were stored at -70°C. Fresh cultures were prepared using a scraping from the surface of an unthawed aliquot to inoculate a flask

TABLE II. *Escherichia coli* strains used.

STRAIN	GENOTYPE	REFERENCE
HB101	<i>F⁻, hsdS20, recA13, ara-14, proA2, lacY1, galK2, strA, xyl-5, mlh-1, supE44</i>	Miller, 1987.
JM101	<i>supE, thi⁻, Δ(lac-pro), F', traD36, proAB, lacI^q ZΔM15</i>	Yanisch-Perron et al., 1985.
JM109	<i>recA1, endoA1, gyrA96, thi⁻, hsdR17 supE44, relA1 Δ(lac, pro) F', traD36, proAB, lacI^q ZΔM15</i>	Miller, 1987.
K38	<i>HfrC(λ)</i>	Russel and Model, 1984.
Y1088	<i>ΔlacU169, supE, supF, hsdR⁻, hsdM⁺, metB, trpR, tonA21, proC::Tn5, (pMC9), pMC9 = pBF.322-lacI^q</i>	Young and Davis, 1983b.
Y1090	<i>ΔlacU169, proA⁺, Δlon, araD139, strA, hflA150[chr::Tn10], (pMC9)</i>	Young and Davis, 1983b.

containing L-medium. HB101 and K38 harbouring recombinant plasmids were also stored in this manner, with the single modification that the original overnight culture was grown in L-medium containing a 1X concentration of the appropriate antibiotic.

B. Description of cloning vectors used

1. Plasmids

The plasmids used were pBR322 which is 4363 base pairs (bp) in size and encodes resistance to both Ampicillin (Amp^r) and tetracycline (Tet^r) (Bolivar et al., 1977; Sutcliffe, 1978 and 1979; Balbas *et al.*, 1986); its derivatives pUC 18 and 19 which are 2686 bp in size and encode Ampicillin resistance (Amp^r) and the alpha peptide of β -galactosidase (Yanisch-Perron et al., 1985); pGEM-2 which is 2869 bp in size, encodes Ampicillin resistance (Amp^r), and possesses both the Bacteriophage T7 and SP6 RNA Polymerase promoters (Promega Biotech); and pGP1-2 which is 7140 bp in size and encodes Kanamycin resistance (Kan^r), Bacteriophage T7 RNA Polymerase, and the temperature sensitive Bacteriophage Lambda repressor cI-857 (Tabor and Richardson, 1985).

2. Bacteriophages

The bacteriophage vectors used were the non-lysogenic phages M13mp18 and M13mp19 (Yanisch-Perron et al., 1985), and the lysogenic bacteriophage lambda gt11 (Young and Davis, 1983a). The M13 vectors differ from each other only in the orientation of the multiple cloning site. In the mature packaged phage particle both of these vectors contain a circular single-stranded (ss) DNA molecule 7249 bp in length encoding the alpha peptide of β -galactosidase. Lambda gt11 contains a linear double stranded (ds) DNA genome 43.7 kbp long encoding the entire *lacZ* gene as well as a temperature sensitive repressor (cI857) which is inactive at 42°C , and an amber mutation (S100) which renders it lysis-defective in hosts which lack the amber suppressor *supF*. A unique EcoRI site is located 53 bp upstream from the β -galactosidase translation termination codon.

3. Recombinant clones

The various recombinant bacteriophage and plasmid clones constructed in the course of this thesis are summarized in Table III.

C. Propagation of bacteriophages

1. M13

a) Preparation of phage containing supernatant

Desired plaques were selected by taking the entire plaque plus the associated agar "plug" in the end of a sterile Pasteur pipette, then expelling the "plug" into a sterile 13x100mm culture tube. The appropriate growth media (generally 1.5 ml of 2X YT) was added to the tube, and the culture incubated at 37°C with agitation at 275 rpm in a New Brunswick Scientific Model G24 Environmental Incubator Shaker for 8 to 12 h. Cell-free, phage containing supernatants were prepared by transferring the entire culture to a 1.5 ml Eppendorf tube and microfuging at 13,000 rpm for 15 min at 4°C. The ensuing supernatant, which contains approximately 10^{13} plaque forming units (pfu) per ml, was transferred to a fresh tube and stored at 4°C until needed.

b) Phage growth and plating

Bacteriophage were plated by adding the desired amount of cell-free phage containing solution to 0.2 ml of a late log phase culture of the appropriate bacterial strain (generally JM101) and 3.0 ml of 50°C top agarose (Appendix A). Bacteriophage, cells, and agarose were mixed by inversion and poured onto a B plate (Appendix A). The plate was incubated in an inverted position overnight (16 to 20 h) at 37°C.

c) Serial dilutions

The cell-free phage containing supernatant from section II.C.1.a. was serially diluted, using TE Buffer (10 mM Tris·Cl pH 7.6, 1 mM Na₂EDTA) as diluant, to 10^{-3} , 10^{-6} , and 10^{-8} times the original titer for replating purposes. Mixing of the phage with the diluant involved vortexing for 5 seconds at the maximum setting on a Fisher Scientific Vortex-Genie.

Table III. Recombinant plasmid and phage clones used.

Name	Vector	Insert Size(kbp)	Restriction Endonuclease Site	Use
L10	pBR322	11.8	BamHI	BamHI TIV DNA containing the L96 gene.
W1	pUC 19	5.5	BamHI/SalI	Probe for 3' end in extended restriction site map.
W3	M13mp19	3.0	EcoRI/SalI	Parental clones used for the generation of Dale deletion subclones for DNA sequencing purposes.
W4	M13mp19	3.0	EcoRI/SalI	
W5	M13mp19	2.5	HindIII	
W6	M13mp19	2.5	HindIII	
W7	pUC 19	3.25	BamHI*/HindIII	Source of the major portion of the L96 ORF.
W8	pUC 19	5.75	EcoRI/HindIII	L96 gene under the control of its native promoter.
W10	pUC 19	3.85	EcoRI/HindIII*	5' fusion of the L96 ORF to the 5' end of the <i>lacZ</i> gene.
W16	pGEM-2	3.85	BamHI/EcoRI	L96 ORF fused to the T7 RNA Polymerase promoter.
W17	pGEM-2	3.85	BamHI	W16 with PTIS insert between promoter and ORF.
W18	pUC 19	0.86	EcoRI/SmaI*	"Sandwich" fusion of L96 C-terminal peptide to <i>lacZ</i> .
W19	pUC 19	3.85	EcoRI/SmaI*	"Sandwich" fusion of complete L96 ORF to <i>lacZ</i> .

* Denotes a vector restriction site not regenerated in the subcloning process.

Generally 10 to 20 pfu were obtained if 10 μ l of the 10^{-8} dilution stock was plated on a petri dish as described in section II.C.1.b.

2. Lambda gt11

a.) Preparation of phage stock

Desired plaques were selected by taking the entire plaque plus the associated agar "plug" in the end of a sterile Pasteur pipette, then expelling the "plug" into 1 ml of Phage Storage Buffer (PSB, 10 mM Tris·Cl pH 7.4, 100 mM NaCl, 10 mM MgCl_2) containing 1 drop of chloroform. The resulting solution was vortexed at the maximum setting on a Fisher Scientific Vortex Genie for 5 seconds, and then stored at 4°C until needed. The titer of this stock was approximately 10^6 pfu/ml.

b) Phage growth and plating

Bacteriophage were plated by adding the desired amount of phage stock (section II.C.2.a.) to 0.2 ml of NZY medium (Appendix A) and 0.1 ml of either Y1088 and Y1090 cells prepared as described below, and incubating at 37°C with agitation at 150 rpm for 15 minutes in a New Brunswick Scientific Model G24 Environmental Incubator Shaker. Warm (50°C) NZY Top Agarose (3.0 ml, Appendix A) was then added and the bacteriophages, cells, and agarose mixed by inversion prior to pouring onto an NZY plate. The resulting plate was incubated in an inverted position overnight (12 to 16 h) at 37°C . The bacterial cells were prepared for plating by inoculating 10 ml of NZY medium with a 1 ml aliquot of an overnight culture of either Y1088 or Y1090 cells. After the incubation of this culture at 37°C with agitation at 275 rpm for 2 h, the cells were harvested by centrifugation at 3000 rpm for 5 min in a Beckman JA-20 rotor at 4°C . Finally, the cells were resuspended in 3 ml of NZY medium by vortexing at a setting of 5, and 3 ml of 10 mM CaCl_2 , 10 mM MgCl_2 solution added to the resulting cell suspension.

c) Serial dilutions

The phage stock from section II.C.2.a. was serially diluted as described for M13 phage (section II.C.1.c.) using Phage Storage Buffer as diluant.

d) *In vitro* packaging of lambda gt11 DNA

In vitro packaging extracts were prepared and utilized to package recombinant lambda gt11 DNAs by Mr. M. Dove (NRC) using the methods of Hohn (1979).

D. Bacterial transformation

Competent cells of the various *Escherichia coli* strains used were prepared and transformed as described by Hannahan (1985). Briefly, this process involved three steps: 1) growth of the cells; 2) preparation of the competent cells; and 3) transformation of these competent cells with exogenous DNA.

A starter culture was prepared by using a loopful of cells of the desired strain (from a single colony on a storage plate prepared as described in section II.A.2.) to inoculate a 25 ml Erlenmeyer flask containing 10 ml of rich medium (usually 2X YT or SOB, see Appendix A). The resulting culture was incubated overnight at 37°C with agitation at 275 rpm in a New Brunswick Scientific Model G24 Environmental Incubator Shaker. A 100 ul aliquot of this culture was then used to inoculate a 250 ml Erlenmeyer flask containing 25 ml of SOB medium. This flask was incubated as described above for 2.5 to 3.0 h until the A_{600} was in the range 0.3 to 0.6. The culture was divided in two equal aliquots and chilled on ice for 15 min before pelleting the cells at 3000 rpm for 12 min in a Beckman JA-20 rotor at 4°C. With the exception of the centrifugation steps, which were done at 4°C, the cells were kept continuously on ice unless otherwise stated. The supernatants were discarded and the two pellets were each resuspended in 4 ml of ice-cold TFB (100 mM KCl, 45 mM $MnCl_2 \cdot 4H_2O$, 10 mM $CaCl_2$, 3 mM Hexamine Cobalt(III) Chloride, and 10 mM 2[N-morpholino]ethane sulphonic acid buffered to pH 6.3 with concentrated KOH) and incubated on ice for 15 min. The cells were again pelleted and the supernatants discarded. After resuspension of each of the cell pellets in 1 ml of ice-cold TFB, the two suspensions were combined in one tube. DMSO (70 ul) was added to the cells with gentle mixing, and the cell suspension incubated on ice for 15 min. To this mixture 70 ul of Dithiothreitol solution (2.2 M Dithiothreitol, 10 mM KOAc) were added with gentle mixing and the cell suspension incubated on ice for 10 min.

Finally, DMSO (70 μ l) was again added to the cell suspension with gentle mixing followed by incubation on ice for 30 min to complete preparation of the competent cells.

Transformation involved placing 210 μ l aliquots of these competent cells in chilled polystyrene culture tubes followed by addition of the desired amount of DNA (generally less than 10 ng) in up to 10 μ l of solution. After gentle mixing the cells were incubated on ice for 30 min. Following this incubation the tubes were heat-shocked in a 42°C water bath for 90 seconds. Subsequent steps differed depending on whether the transforming DNA was a plasmid or viral in origin. If plasmid DNA was used the heat shocked tubes were incubated on ice for 2 min before adding 800 μ l of SOC medium (Appendix A) to each and incubating at 275 rpm at 37°C for 1 h. These cultures were plated by mixing the appropriate amount of the transformed *E. coli* culture with 10 μ l of IPTG solution (Appendix A), 50 μ l of Bluo-Gal solution (Appendix A), if desired, and sufficient water to yield a final volume of 260 μ l. The resulting solution was spread evenly over the desired type of petri dish and allowed to absorb into the agar. If bacteriophage replicating form (RF) DNA was used the heat shocked tubes had IPTG and Bluo-Gal added to them as in plasmid transformations. This was followed by the addition of 0.2 ml of "fresh" cells (i.e. taken from a late log phase culture of the particular *E. coli* strain being transformed) to serve as a bacterial "lawn" for virus propagation, and 3 ml of 50⁰ top agarose (Appendix A). The resulting solution was mixed by inverting three times and then evenly poured over the desired type of petri dish and allowed to harden. In both types of transformations the petri dishes were incubated in an inverted position overnight at 37°C.

E. Isolation and characterization of plasmid and phage DNA

Throughout the following protocols the water used was distilled, de-ionized, and sterilized by filtration through a 0.22 μ m filter.

1. Plasmids

a) Purification of DNA fragments via extraction with organic solvents and ethanol precipitation

Generally phenol, phenol/chloroform, and chloroform, all prepared as described by Maniatis *et al.* (1982), were used for extraction purposes. If the volume of the solution to be extracted was less than 100 μ l sufficient water was added to bring the final volume up to this level prior to performing the extraction step. The solutions to be extracted were mixed, either by inversion or by vortexing at a setting of 3 on a Fisher Scientific Vortex-Genie, with an equal volume of the phenol, phenol/chloroform, or chloroform. Separation of the aqueous (top) and organic (bottom) phases was achieved by centrifuging at 13,000 rpm in a microfuge at room temperature for 5 min. After centrifugation the aqueous phase was carefully transferred to a fresh tube. The DNA was precipitated by adding one-half volume of 7.5 M NH_4OAc , mixing, adding two volumes of cold (-20°C) ethanol, again mixing, and incubating either at -70°C for 30 min or at -20°C overnight. The DNA was pelleted by microfuging at 4°C for 30 min at 13,000 rpm. After removal of the supernatant, the pellet was washed by adding cold (-20°C) 70% ethanol to two-thirds of the final precipitating solution volume, mixing, and microfuging as above for 15 min. The ensuing pellet was dried under vacuum without heating in a Savant Speed Vac sample concentrator for 15 min. The pellet was dissolved in water to the approximate desired concentration, and the concentration verified by agarose gel electrophoresis (see section II.E.1.e.).

b) Extraction of DNA fragments from agarose gels (electroelution)

Electroelutions were performed according to the method of Dretzen *et al.* (1981) using NA45 membrane (Schleicher & Schuell). The membrane was prepared by cutting it to the desired size and soaking it in 2.5 M NaCl for several hours followed by several washes in distilled water, and was stored until needed in 1 mM Na_2EDTA pH 8 solution at 4°C . The sample to be electroeluted was electrophoresed as described in section II.E.1.e. When the electrophoresis was complete the gel was examined under long wavelength ultraviolet light

(366nm) and a slit made in the gel directly in front of the band to be electroeluted. The prepared NA45 membrane was inserted into this slit and the electrophoresis resumed. When all of the DNA in the desired band had migrated onto the membrane, as verified by examination under long wavelength ultraviolet light, the membrane was removed from the gel, washed with water, and placed in an Eppendorf tube containing 300 μ l of "elution solution" (2.5 M NaOAc, 20 mM Tris·Cl pH 7.6, 1 mM Na₂EDTA). The tube was incubated at 65°C for 1 h, and the resulting solution transferred to a fresh tube. The original tube was washed twice with 100 μ l of "elution solution", and these washings were added to the original solution to give a final volume of 500 μ l. This solution was extracted once with phenol/chloroform and ethanol precipitated as described in section II.E.1.a. with the single modification that no NH₄OAc was added.

c) Removal of contaminants from DNA by spin columns

Spin columns (Maniatis et al., 1982) were prepared using a suspension of Sephadex G50 in TE buffer. The column was prepared by adding this suspension to a needle-less 1 ml syringe barrel in which a small plug of sterile glass wool had been placed. The TE buffer was run out of the column by centrifuging at 2000 rpm in a Beckman Accuspin Fr centrifuge for 1 min. This loading procedure was repeated until the level of Sephadex in the syringe barrel after centrifuging reached the 0.9 ml graduation on the barrel. The column was then washed twice by loading 200 μ l of STE buffer (10 mM Tris·Cl pH 8, 1 mM Na₂EDTA, 100 mM NaCl) on the column and running it through by centrifuging for 1 min at 2000 rpm. The sample, dissolved in STE buffer, was then loaded on the column and the column centrifuged for 1 min at 2000 rpm. The eluate was collected in an Eppendorf tube and the DNA ethanol precipitated by adding two volumes cold (-20°C) ethanol directly to the column eluate and treating as described in section II.E.1.a.

d) Small scale plasmid preparation

Small scale (20-40 μ g) plasmid preparations were done using the Alkaline Lysis Method of Maniatis et al. (1982). In this protocol isolated colonies were used to inoculate 5 ml cultures

in rich medium (usually L Medium, see Appendix A) containing the appropriate amount of the relevant antibiotic. The cultures were grown overnight at 37°C with agitation at 275 rpm in a New Brunswick Scientific Model G24 Environmental Incubator Shaker. A 1.5 ml aliquot of this culture was used for preparation of the plasmid DNA, while the bulk of the liquid culture was stored at 4°C. This 1.5 ml aliquot was microfuged at 13,000 rpm at 4°C for 1.5 min. The supernatant was discarded and the open tube placed in an inverted position on a paper towel for 15 min. The pellet was resuspended by vortexing at a setting of 3 to 5 on a Fisher Scientific Vortex-Genie in 100 ul of ice-cold 50 mM glucose, 10 mM Na₂EDTA, 25 mM Tris·Cl pH 8 solution which contained 4 mg/ml lysozyme (powdered lysozyme was added to the solution just prior to use). The solution was incubated at room temperature for 5 min before adding 200 ul of 0.2 M NaOH, 1% SDS solution to it and mixing by inverting five times. This solution was incubated on ice for 5 min before adding 150 ul of ice-cold potassium acetate solution (prepared as described by Maniatis *et al.*, 1982), gently vortexing at a setting of 2 to 3 in an inverted position for 10 seconds, and incubating on ice for 5 min before microfuging at 13,000 rpm at 4°C for 5 min. The supernatant was transferred to a fresh tube and extracted once with one volume of phenol/chloroform. After microfuging to separate the phases, the aqueous phase was transferred to a fresh tube and the DNA precipitated by adding two volumes of room temperature ethanol and vortexing gently at a setting of 2 to 3 for 5 seconds. This solution was incubated at room temperature for 2 min before microfuging at 13,000 rpm at room temperature for 5 min. The pellet was washed by adding 1 ml of room temperature 70% ethanol and mixing by vortexing very gently at a setting of 2 for 2 seconds. The pellet was collected as described above and drained of ethanol by placing the centrifuge tube in an inverted position on a paper towel for 15 min. The periphery of the tube was wiped with a cotton swab and the pellet dried under vacuum as described in section II.E.1.a. The dried pellet was dissolved in 50 ul of TE buffer containing 20 ug/ml RNase A (prepared as described by Maniatis *et al.*, 1982) and the resulting solution stored at 4°C.

e) Restriction endonuclease digestion and agarose gel electrophoresis

Analytical restriction endonuclease digests were performed by adding 0.5-1.0 μ g of the DNA to be analyzed to a solution containing 1.0 μ l of a 10X stock of the appropriate restriction endonuclease buffer (Maniatis *et al.*, 1982), 10 units of the restriction endonuclease of interest, and sufficient water such that the final reaction volume was 10 μ l. The digest was incubated at the manufacturer's recommended temperature for 1 to 2 h before adding 3 μ l of 5X STOP Buffer (100 mM Na₂EDTA, 0.5% SDS, 25% glycerol, 0.1% Bromophenol Blue) to the sample and electrophoresing in an appropriate percentage agarose (SeaKem ME Agarose) slab gel. Restriction endonucleases were purchased from Bethesda Research Laboratories, International Biotechnologies Inc., New England Biolabs, and Pharmacia, and were used interchangeably.

Agarose gels were prepared by heating to dissolve the appropriate amount of solid agarose (0.7-1.5% final) in the desired volume of either 1X TAE buffer (Maniatis *et al.*, 1982) for dsDNA gels, or 1X TBE buffer (Biggin *et al.*, 1983) for ssDNA gels (see Appendix A for the preparation of TAE and TBE Buffers). For dsDNA gels ethidium bromide (EtBr, 0.8 μ g/ml final) was added just prior to pouring the gels. For ssDNA the gels were run in the absence of EtBr and were stained after electrophoresis by soaking in EtBr solution (5 μ g/ml) for 5 min, followed by destaining for 20 min in 1X TBE buffer. Both types of gels were photographed through a red filter under ultraviolet light (254nm) using Polaroid Type 52 or 55 Polapan Film in a Polaroid Land Camera.

Both HindIII digested and EcoRI/HindIII codigested bacteriophage Lambda DNA were used as molecular weight markers in agarose gels. The former was taken as yielding fragments 23.130, 9.419, 6.557, 4.371, 2.322, 2.028, 0.564, and 0.125 kbp in length, while the latter was taken as yielding fragments 21.7, 5.15, 5.0, 4.27, 3.48, 1.98, 1.9, 1.59, 1.37, 0.940, 0.830, 0.560, and 0.140 kbp in length (1985 IBI catalogue).

f) 5' and 3' DNA terminal modifications

(1) 5' overhangs

5' overhangs, or recessed 3' ends, were filled in using the Klenow fragment of DNA polymerase I (Maniatis *et al.*, 1982). The reaction involved adding 1 ug of the DNA fragment to a solution containing 2.0 ul of 10X Nick Translation Buffer (500 mM Tris·Cl pH 7.2, 100 mM MgSO₄, 1 mM Dithiothreitol (Calbiochem), 500 ug/ml Bovine Serum Albumin (BRL)), 1.0 ul of a 10 mM solution of each of the required dNTPs (prepared as described by Maniatis *et al.* (1982)), 1 unit of the Klenow fragment of DNA polymerase I, and sufficient water to yield a final volume of 20 ul. The resulting solution was incubated at room temperature for 30 min before heat inactivating the enzyme at 65°C for 10 min. The reaction solution was then extracted once with phenol/chloroform and the DNA ethanol precipitated (see section II.E.1.a).

Alternatively, 5' overhangs were removed using Mung Bean Nuclease via the protocol described by Maniatis *et al.* (1982). The DNA fragment of interest was added to a solution containing 1.5 ul of 10X Mung Bean Nuclease Buffer (300 mM NaOAc pH 4.6, 500 mM NaCl, 10 mM ZnCl₂, 50% glycerol), 3 units Mung Bean Nuclease, and sufficient water to yield a final volume of 15 ul. The reaction was incubated at 37°C for 10 min, and was quenched by adding 5 ul of 250 mM Na₂EDTA pH 8.0 solution. The resulting solution was phenol/chloroform extracted and ethanol precipitated as described in section II.E.1.a.

(2) 5' dephosphorylation

Terminal 5' phosphates were removed from fragments using Calf Intestinal Alkaline Phosphatase (CIP) as described by Maniatis *et al.* (1982). This is a desirable step in any cloning experiment as a dephosphorylated vector is incapable of ligating to itself; hence, any clones which appear should be recombinant. The fragment to be dephosphorylated (0.5-1 ug) was added to a solution containing 5 ul of 10X CIP buffer (500 mM Tris·Cl pH 9, 10 mM MgCl₂, 1 mM ZnCl₂, 10 mM spermidine), 1.0 ul of CIP (as a suspension, this usually is 15 units), and sufficient water to yield a final volume of 49 ul. The reaction was incubated at

37°C for 30 min. A second 1.0 ul aliquot of the enzyme was then added and the incubation at 37°C continued for 30 min. The reaction was quenched by adding 40 ul of water, 10 ul of 10X STE buffer, and 5 ul of 10% SDS. The solution was mixed and incubated at 68°C for 15 min before extracting twice with both phenol/chloroform and chloroform (see section II.E.1.a.). The final aqueous phase was passed through a spin column and the DNA ethanol precipitated (see section II.E.1.c.).

g) DNA ligation

Ligations were performed by mixing the vector (20-50 ng) and insert DNAs in a 1:3 molar ratio and adding 1.0 ul of 100 mM Dithiothreitol, 1.0 ul of 250 mM Tris·Cl pH 7.6, 100 mM MgCl₂ buffer solution, either 1.0 ul (sticky end ligations) or 0.5 ul (blunt end ligations) of 10 mM ATP, 6-8 units of T4 DNA Ligase, and sufficient water to yield a final volume of 10 ul. The ligation was incubated overnight either at 16°C for sticky end ligations, or at room temperature for blunt end ligations. A fraction of the ligation mixture was used to transform a suitable strain of *E. coli* (section II.D.), while the remainder was stored at 4°C until needed.

h) Radiolabelling of DNA and hybridization

(1) Nick translation

Purified DNA fragments were radiolabelled by nick translation (Maniatis *et al.*, 1982). The fragment to be labelled (200-500 ng) was added to a solution containing 5 ul of 10X Nick Translation Buffer (see section II.E.1.f.), 3 ul of a 1 mM solution of all dNTPs except dATP, 50-100 uCi of [α -³²P]dATP, sufficient water to yield a final volume of 48 ul, and 1 ul of a 10⁻⁴ diluted solution of DNase I. This solution was incubated at room temperature for 3 to 5 min before adding 10 units of DNA Polymerase I to the reaction and incubating at 16°C for 1 h. The reaction was stopped by the addition of 2 ul of 250 mM Na₂EDTA pH 8 and 48 ul of STE Buffer, and the resulting solution was passed successively through two spin columns (see section II.E.1.c.) to yield the final solution. The level of radioactivity of the resulting probe was determined by adding 1 ul of this solution to approximately 10 ml scintillation fluid (prepared by mixing 5 parts 0.06% 2,5-diphenyloxazole (New England Nuclear) in toluene

with 1 part ethanol) in a glass scintillation vial, and scintillation counting in a Beckman Model LS 7000 scintillation counter. Approximately $0.5-1 \times 10^6$ cpm of probe were used per ml of hybridization fluid for the probing of homologous DNA.

(2) Colony hybridization

Colony hybridizations were performed using the procedure of Maniatis *et al.* (1982) with modifications. Biodyne nylon membrane was cut to the desired shape, labelled, and marked so that the orientation could be determined after the autoradiography. The filter was placed on the petri dish harboring the colonies to be screened for 1 to 2 min, then removed and placed cell side up on approximately 250 ul of 0.3 M NaOH, 1.5 M NaCl solution on a piece of plastic wrap for 5 min. The filter was transferred (cell side up) to 250 ul of 0.5 M Tris·Cl pH 8, 3 M NaCl solution on a piece of plastic wrap for 3 min. The filter was then transferred (cell side up) to 2X SSC solution (300 mM NaCl, 30 mM sodium citrate pH 7) for 5 min before being placed between two pieces of Whatman 3MM filter paper and being baked at 80°C for 30 min. The dried filter was placed in 2X SSC, 0.1% SDS solution at room temperature for 2 to 24 h. Any colony remnants which were left on the filter at this point were gently wiped away with a finger of a gloved hand. The filter was air dried or used directly in the prehybridization step.

The filters were prehybridized by placing them in a hybridization sac along with 30-50 ml of preheated (65°C) 1 M NaCl, 1% SDS, 100 ug/ml salmon sperm DNA solution. The sac was sealed and incubated in a 65°C water bath for one hour. During this time the radiolabelled probe was boiled for 10 min. The prehybridization solution was removed from the sac and replaced with 10-15 ml of fresh, preheated (65°C) solution. The boiled probe was added to this solution, the sac sealed, and then incubated overnight in a 65°C water bath. The filters were washed thrice for 5 min each with shaking in room temperature 2X SSC, 0.1%SDS solution; twice for 30 min without shaking at 65°C in 2X SSC, 0.1% SDS solution; and once for 10 min at room temperature without shaking in 1X SSC solution. While still damp the filters were wrapped in plastic wrap and placed in an X-ray cassette along with an

unexposed Kodak X-Omat RP film. After 1 to 3 days the film was developed as prescribed by the manufacturer.

(3) Southern blotting

The transfer of dsDNA from agarose slab gels (see section II.E.1.e.) to nylon membranes was performed using the alkaline transfer method of Reed and Mann (1985). Prior to the transfer of the DNA to the membrane the gel was incubated with gentle rocking at room temperature in 0.25 N HCl for 10 min. This "acid nicking" step depurinates the DNA and facilitates the transfer of large fragments to the membrane. The blot was mounted by cutting two wicks from 3MM Whatman filter paper to the exact width of the gel. These wicks were wet in 0.4 N NaOH and placed on a plastic support (i.e. an inverted gel casting tray) in a larger reservoir container such that the ends of the wicks were immersed in 0.4 N NaOH solution. The wells were then cut off and the gel placed on top of the wicks taking care that no bubbles were trapped between the gel and the wicks. Precut Hybond-N (Amersham) nylon membrane was wet in water and placed on top of the gel, again taking care to ensure that no bubbles were trapped. Mounting of the blot was completed by placing 6 pieces of 3MM Whatman filter paper, 20 pieces of blotting paper (Bethesda Research Laboratories), approximately one inch of Scott paper towels, and an appropriately sized weight on top of the nylon membrane. The transfer was allowed to proceed for 6 h at which point the blot was dismantled and the nylon membrane incubated in 2X SSC solution for 10 min at room temperature to neutralize the alkali. The DNA was fixed to the membrane by baking at 80°C for 1 h. The filter was then prehybridized, hybridized, and washed as described in section II.E.a.h.2. with the exception that all washes were performed with gentle agitation at the appropriate temperature. The washed filter was wrapped in plastic wrap and autoradiographed in the usual manner.

2. Bacteriophages

a) M13

(1) Extraction and purification

(a) RF DNA preparation

Double stranded M13 replicating form DNA was isolated using the Alkaline Lysis Method of Maniatis *et al.* (1982) for preparing small (20-40 ug) amounts of plasmid DNA (see section II.E.1.d.).

(b) Viral "+" strand DNA Preparation

Both large and small scale M13 "+" strand viral DNA isolation protocols were used (Messing, 1983). The large scale preparation required an overnight culture of *E. coli* strain JM101 (Table I) to be grown in 2X YT medium (Appendix A). A 100 ul aliquot of this culture was used to inoculate 10 ml of 2X YT medium in a 25 ml Erlenmeyer flask. The resulting culture was incubated at 37°C with agitation at 275 rpm at 37°C in a New Brunswick Scientific Model G24 Environmental Incubator Shaker for 30 min. The cells were infected by adding 50 ul of the approximately 10^{12} pfu/ml cell-free supernatant containing the phage whose "+" strand DNA is to be isolated (prepared as described in section II.C.1.a.) to the culture and the incubation continued for 8 to 16 h. The culture was transferred to a Corex tube and centrifuged for 5 min at 10,000 rpm in a Beckman JA-20 rotor at 4°C. The supernatant was transferred to a fresh Corex tube and centrifuged as described above for 15 min. Again the supernatant was transferred to a fresh Corex tube, and its volume was determined. For every 10 ml of supernatant volume, 2 ml of 20% w/v PEG-6000, 2.5 M NaOAc pH 7 solution were added. The tube was covered with parafilm and the contents mixed by inverting. The resulting solution was incubated at 4°C for 30 min before centrifuging for 20 min at 10,000 rpm at 4°C. The supernatant was discarded and the tube left in an inverted position on a paper towel for 1 h. At this point any excess remaining solution which could be wiped away with a cotton swab without disturbing the phage pellet was removed. The phage pellet was resuspended in 500 ul of STE Buffer for every 10 ml of

original supernatant volume, and the suspension transferred to an Eppendorf tube and extracted once with phenol, twice with phenol/chloroform, and once with chloroform before ethanol precipitating the DNA as described in section II.E.1.a. with the single modification that no NH_4OAc was added to the solution prior to the precipitation step. The dried pellet was dissolved in 20 μl of TE^* Buffer (10 mM Tris·Cl pH 7.6, 0.1 mM Na_2EDTA). The optical density of 1 μl of the resulting solution at 260 nm was determined and the concentration of the DNA estimated by comparison to the standard of 1 OD_{260} representing a concentration of 33 $\mu\text{g}/\text{ml}$. The concentration of the purified "+" strand DNA was adjusted to 0.5 $\mu\text{g}/\mu\text{l}$ by adding the appropriate volume of TE^* Buffer, and the solution stored at 4°C. A 1.0 μl aliquot of this solution was analyzed by adding 6 μl of water and 3 μl of 5X STOP Buffer to it and electrophoresing the resulting solution as described in section II.E.1.e.

Small scale M13 "+" strand viral DNA preparations were performed using a similar protocol. A 1.0 ml aliquot of the approximately 10^{13} pfu/ml cell-free supernatant containing the phage whose "+" strand DNA is to be isolated was placed in an Eppendorf tube, and 200 μl of 20 % w/v PEG-6000, 2.5 M NaOAc pH 7 solution added. The contents of the tube were mixed by inverting 50 times before incubating at 4°C for 30 min. The solution was microfuged at 13,000 rpm at 4°C for 15 min, and the supernatant discarded. The open tube was left in an inverted position on a paper towel for 30 min after which the periphery of the tube was wiped with a cotton swab. The pellet was resuspended in 600 μl of STE buffer by vortexing for 5 seconds. The suspension was extracted once with phenol, once with phenol/chloroform, and once with chloroform. The DNA in the final aqueous phase was ethanol precipitated in a manner identical to that in the large scale preparation. The dried pellet was dissolved in 15 μl of TE^* Buffer and stored at 4°C. A 1 μl aliquot of this solution was analyzed by agarose gel electrophoresis as described above in the large scale preparation.

(2) Agarose gel electrophoresis

(a) Phage DNA

Recombinant bacteriophage were analyzed on agarose gels by first lysing the phage, and then electrophoresing the lysate on a ssDNA gel (Messing, 1983) prepared as described in section II.E.1.e. The lysates were prepared by mixing 20 μ l of the approximately 10^{13} pfu/ml phage containing supernatant from section II.C.1.a with 5 μ l of Phage Lysis Buffer. The solution was mixed by gently tapping the tube, and then incubated at room temperature for 1 h before electrophoresing a 10-20 μ l aliquot. Phage Lysis Buffer was prepared by mixing 900 μ l of 5X STOP Buffer with 100 μ l of 20% SDS.

(b) C-test

Because the "+" strand of the viral DNA is always packaged into the M13 viral particle, two recombinant phage with a DNA fragment cloned in opposite orientations into the RF will hybridize only via the insert region. This fact was used to test whether two recombinant phages had DNA complementary to each other using the protocol of Messing (1983), with modifications. Briefly, the C-test was performed by mixing 10 μ l of the approximately 10^{13} pfu/ml phage containing supernatant (see section II.C.1.a.) of each of the two recombinant phages of interest with 1 μ l of 2.5 M NaCl and 5 μ l of C-Buffer (60% de-ionized formamide, 1% SDS, 0.1% Bromophenol Blue). The sample, along with similarly prepared positive and negative controls, was incubated at 68°C for 1 h before electrophoresing it through a 0.7% agarose gel in 1X TBE Buffer (Biggin *et al.*, 1983) containing 0.8 μ g/ml EtBr. If the two viral DNAs hybridize via their insert region, they form a figure eight-like structure and migrate slower in the gel than the single-stranded viral circles. Thus, the presence in the gel of bands higher than either of the viral bands indicates that the two phages being analyzed have complementary inserts.

(c) Insert sizing using Nuclease S1

The formation of a figure eight-like structure when two recombinant phages harbouring complementary inserts are hybridized together was also used to accurately estimate the size

of the smaller of the two inserts (Baird, 1989). The "+" strand DNAs of the two phages to be tested for complementary inserts were annealed by mixing 1 ul of each (prepared as described in section II.E.2.a.1.b.) with 6 ul of water and 1 ul of 10X Nuclease S1 Buffer (2 M NaCl, 0.5 M NaOAc pH 4.5, 10 mM ZnSO₄, 5% glycerol) in an Eppendorf tube, and incubating at 65°C for 20 min. The single-stranded regions of the figure eight-like structure were then digested by adding 1 ul of 1 U/ul Nuclease S1 solution and incubating at 37°C for 20 min. The reaction was quenched by the addition 3 ul of 5X STOP Buffer and electrophoresed as described in section II.E.1.e for dsDNA. Since Nuclease S1 degrades ssDNA, but not dsDNA, the figure eight-like structure is digested by the enzyme leaving only the linear double stranded section created by the complementary inserts of the two phages. Electrophoresis of this fragment and comparison to markers yields an accurate estimation of the insert size of the smaller of the two recombinant phages.

(3) Cloning and subcloning

All cloning and subcloning in M13 vectors were done by considering the RF as being a plasmid. All manipulations of this DNA were performed using the protocols described for plasmid DNA in section II.E.1.

(4) Construction of deletion clones for DNA sequencing

Deletion subclones of M13 parental clones for DNA sequencing purposes were constructed using the method of Dale *et al.* (1985). The protocol involved mixing 5 ug of purified "+" strand DNA (see section II.E.2.a.1.b.) of the parental clone to be deleted with 2.5 ul of 10X Dale Deletion Buffer (330 mM Tris base, 660 mM KOAc, 100 mM Mg(OAc)₂, with the pH of the complete solution adjusted to 7.8 with HOAc), 1 ul of 100 mM Dithiothreitol, 4 ul of 0.5 A₂₆₀/ml RD20 primer (5'-CGACGGCCAGTGAATTCCCC-3'), and sufficient water to yield a final volume of 22.5 ul. This solution was placed in a 65°C water bath for 10 min, then the water bath was removed from the heat source and allowed to slowly cool to 35°C thereby annealing the primer to the viral DNA. At this point the tube was removed from the bath, 2 ul of 50 U/ul EcoRI solution added to it, and the digest

incubated at 50°C for 2 to 16 h. A 1 ul aliquot of the digest was removed and added to an Eppendorf tube containing 5 ul of water and 2 ul of 5X STOP Buffer. The resulting solution, along with an undigested control, was electrophoresed on a 0.7% agarose gel in 1X TBE Buffer (Biggin *et al.*, 1983) containing 0.8 ug/ml EtBr. If the gel revealed that the digestion was incomplete more enzyme was added and the incubation continued. If the gel indicated that the digestion was complete then the enzyme was inactivated by heating at 65°C for 10 min.

Once cool, 2.5 ul of 100 mM Dithiothreitol, 1 ul of 10 mg/ml Bovine Serum Albumin, and 1 ul of 6 U/ul T4 DNA Polymerase (IBI) were added. The deletion reaction was incubated at 37°C and, assuming a digestion rate of approximately 50 bases per minute, appropriately sized aliquots were withdrawn at various time intervals in order that deletion subclones with the desired differences in size were obtained. These aliquots were pooled in an Eppendorf tube stored on dry ice until the last aliquot had been removed. The enzyme was then inactivated as described above.

After cooling to room temperature, the G-tailing reaction was performed by adding 3 ul of 50 uM dGTP, 1 ul of water, and 1 ul of 12 U/ul TdT (Pharmacia or IBI) to the inactivated digestion reaction and incubating at 37°C for 10 min. Following heat inactivation of the TdT at 65°C, a 10 ul aliquot of the solution was set aside as an unligated control.

Recircularization of the DNA in the remainder of the solution involved annealing 1 ul of 0.5 A₂₆₀/ml RD20 primer to the DNA as described above, followed by the addition of 3 ul of 10 mM ATP, 4 ul of water, and 6 to 8 units of T4 DNA Ligase (Pharmacia) and incubating at room temperature for 2 to 4 h.

During this time competent *E. coli* strain JM101 cells were prepared as described in section II.D. Approximately one-half of both the ligated sample (with 2-3 ul aliquots of the ligation reaction being plated on each of 5-6 petri dishes) and the unligated sample (with all 5 ul of this half of the unligated sample being plated on one petri dish) were transfected into these competent cells. The number of plaques on the plates possessing the cells transfected

with the ligated DNA was compared to the number on the plates possessing the cells transfected with the unligated DNA. Ratios of 20 to 50 were desired, anything less than 20 indicated the presence of numerous undeleted parental clones on the ligated sample plates. Plaques on the ligated sample plates were analyzed by preparing phage containing supernatants (see section II.C.1.a.) and electrophoresing phage lysates through agarose gels (see section II.E.2.a.2.a.). From these experiments the clones required to sequence to fragment of interest were selected and "+" strand DNA isolated (see section II.E.2.a.1.b.). The insert size of these selected clones was estimated by Nuclease S1 sizing (see section II.E.2.a.2.c.), and from these results the final clones required for DNA sequencing were selected.

b) Lambda gt11

(1) Extraction and purification

Bacteriophage lambda gt11 DNAs were isolated using the plate lysate method of Maniatis *et al.* (1982). A 50 to 100 μ l aliquot of phage stock from section II.C. 2.a. (approximately 10^5 pfu) was plated on an NZY plate containing 1.5% agarose instead of agar as described in section II.C.2.b. The resulting plate was incubated at 37°C in an inverted position overnight until the plaques covered almost the entire surface of the plate (9 to 14 h). The plate was removed from the incubator, covered with 5 ml of Phage Storage Buffer, and incubated at room temperature for 2 h. The solution was then recovered from the top of the plate and centrifuged for 10 min at 9000 rpm in the Beckman JA-20 rotor at 4°C. The resulting supernatant was transferred to a fresh tube and both RNase A and DNase I were added to a final concentration of 1 μ g/ml each. The solution was then incubated at 37°C for 30 min before adding 1 volume of 20% PEG-6000, 2.5 M NaOAc pH 7 solution. The resulting solution was mixed by inversion and stored on ice for 1 h. The precipitated bacteriophage were recovered by centrifuging for 30 min at 10,000 rpm in the Beckman JA-20 rotor at 4°C. The tubes were left in an inverted position on a paper towel for 30 min before resuspending the pellet in 500 μ l of Phage Storage Buffer by vortexing at the maximum setting on a Fisher

Scientific Vortex Genie. Any remaining debris was removed by centrifugation at 9000 rpm for 2 min and transferring the supernatant to a fresh tube. Both 10% SDS solution (5 ul) and 0.5 M Na₂EDTA solution (5 ul) were added and the resulting solution extracted once each with one volume of phenol, phenol/chloroform, and chloroform as described in section II.E.1.a. The DNA was precipitated by the addition of 1 volume of room temperature isopropanol and mixing by inversion followed by incubation at -70°C for 20 min. The DNA was pelleted at 13,000 rpm for 15 min in a microfuge. The DNA pellet was washed and dried as described in section II.E.1.a., and dissolved in 50 ul of TE Buffer containing 20 ug/ml of RNase A. The resulting solution was stored at 4°C.

(2) Cloning and subcloning

All manipulations of lambda gt11 DNA for cloning and subcloning purposes were performed using the protocols described for plasmid DNA in section II.E.1.

(3) Plaque hybridizations

Plaque hybridizations were performed using the protocol described in section II.E.1.h.2 for colony hybridizations.

F. Isolation and characterization of TIV

TIV infected *Galleria mellonella* larvae were prepared and stored at -20°C as described by Yule and Lee (1973), and were generously provided by Dr. S. Tajbakhsh.

1. Isolation of TIV virions from infected larvae

TIV virions were purified from infected *Galleria mellonella* larvae as described by Yule and Lee (1973). Approximately 80 frozen TIV infected *Galleria mellonella* larvae were placed in a small amount of liquid nitrogen and ground up using a mortar and pestle. The ensuing powder was made into a slurry with 10 mM Sodium Phosphate Buffer pH 7.2 (prepared by mixing 2.8 ml 1 M NaH₂PO₄ with 7.2 ml 1 M Na₂HPO₄ and diluting the resulting solution to one liter with water). The slurry was hand stirred to suspend the virus particles, and then filtered through cheese cloth to remove the larval debris. The virions were purified by a series of alternating low and high speed centrifugation steps. The filtrate was initially

centrifuged at 5000 rpm in a Beckman JA-20 rotor for 5 min, and the supernatant poured through a glass wool plug into fresh tubes. The virions were pelleted by ultracentrifugation at 25,000 rpm for 1 h at 4°C in a Beckman Type 30 rotor. The pellets (4) were resuspended in approximately 5 ml of Sodium Phosphate Buffer each, and subjected to a second low speed centrifugation. Again the resulting supernatants were passed through a glass wool plug into fresh tubes and were subjected to a second high speed centrifugation. The resulting virion pellets (4) were resuspended in 2 ml of Sodium Phosphate Buffer each, and the suspensions pooled in a beaker. The centrifuge tubes were washed by successively passing 1 ml of Sodium Phosphate Buffer through them, and then combining this wash solution with the virion suspension in the beaker to yield a final volume of approximately 10 ml in the beaker.

Approximately 3.3 ml aliquots of this viral suspension were then layered on each of three 20 ml linear 10-40% sucrose gradients and centrifuged for 30 min at 15,000 rpm in the Beckman SW 25.1 swinging bucket rotor at 4°C. The main viral band was recovered from each of the three tubes and placed in fresh tubes. In order to dilute out the sucrose in the samples the tubes were filled with Sodium Phosphate Buffer, mixed by inversion, and the virions pelleted at high speed as described above. The ensuing pellets were resuspended in Sodium Phosphate Buffer, and the density gradient centrifugation step repeated on a 10-45% linear sucrose gradient. Again the main viral bands were removed, pooled in one tube, and the sucrose diluted out by filling the tube with Sodium Phosphate Buffer. The solution was divided into three equal aliquots and subjected to a final high speed spin. The ensuing blue pellets (approximately 1 mg each) were resuspended in 2 ml of Sodium Phosphate Buffer each and stored at 4°C until needed.

2. Isolation of total TIV DNA from purified virions

DNA was purified from TIV virions using the procedure of Tajbakhsh *et al.* (1986) with minor modifications (Tajbakhsh, 1988). Purified TIV virions from the preceding section (1 mg pellet) were resuspended in Sodium Phosphate Buffer (2 ml, see the preceding section), and the volume made up to 4 ml with 20 mM Tris·Cl pH 8, 10 mM MgSO₄, 100 mM

NaCl solution. SDS, Na₂EDTA, and Proteinase K were then added to final concentrations of 2%, 20 mM, and 150 ug/ml, respectively. The solution was mixed by gentle hand inversion and incubated at 42°C for 2 h. The DNA was extracted three times with one volume of phenol/chloroform (see section II.E.1.a) with the phases being separated by centrifugation for 15 min at 3000 rpm in a Beckman JA-20 rotor. The three organic layers were washed by successively extracting them with 1 ml of TE Buffer, and then combining this washing with the final aqueous phase. The resulting solution was dialyzed against 2 liters of STE buffer for 48 h. The solution was recovered from the dialysis bag into a 50 ml polypropylene tube. The inside of the dialysis bag was rinsed with 3 ml of STE Buffer, and this washing combined with the initial solution to yield a final volume of 8.5 ml. The DNA was precipitated by layering on top of the aqueous phase two volumes of room temperature 95% ethanol. The contents of the tube were mixed by gentle hand inversion over a period of 15 to 20 min, and the DNA fibres recovered on a glass rod. The DNA coated tip of the glass rod was dipped in 70% ethanol to desalt the DNA, and then placed in 200 ul of TE Buffer in an 1.5 ml Eppendorf tube and the DNA pellet dislodged from the glass rod. The dissolution of the DNA was aided by incubating the tube at 37°C for 4 h prior to storing it at 4°C for several days. The concentration of the DNA from optical density measurements was found to be 0.44 ug/ul.

G. Analysis of DNA by nucleotide sequencing

Both the Klenow fragment of *E. coli* DNA Polymerase I and the genetically engineered Bacteriophage T7 DNA Polymerase called "Sequenase" (United States Biochemical Corporation) were used in the dideoxy chain termination method of Sanger *et al.* (1977) for DNA sequencing purposes. All sequencing was performed using [α -³⁵S]dATP as radioactive label, and, with two exceptions, only commercially available M13 universal sequencing primers were used. The two exceptions involved custom synthesized oligonucleotide primers (Dr. W. Sung, NRC) which were used to determine sequence from existing clones that could not be elucidated using the commercially available primers.

1. Reactions

a) DNA synthesis using DNA-dependent DNA polymerases

(1) Klenow fragment of *E. coli* DNA polymerase I

DNA Sequencing using the Klenow fragment of *E. coli* DNA Polymerase I (Messing, 1983) involved first annealing the primer to the purified M13 "+" strand DNA, followed by extension of this primer by the enzyme. The primer was annealed by mixing 3.5 μ l of the viral "+" strand DNA (purified as described in section II.E.2.a.1.b.) with 2 μ l of 0.03 A₂₆₀/ml primer, 1.5 μ l of 100 mM Tris·Cl pH8, 50 mM MgCl₂ buffer solution, and sufficient water to yield a final volume of 10 μ l, heating in a 65°C water bath for 10 min, followed by removal of the heat source and gradual cooling of the bath. When the temperature of the water bath reached 35°C the tube was removed from the bath. The polymerase reaction was prepared by adding 2.5 μ l of 10 mCi/ml, 600 Ci/mmol [α -³⁵S]dATP and 1.5 μ l of 1 unit/ μ l Klenow Fragment of DNA Polymerase I to the primer annealed DNA, mixing by gently tapping the tube, and then placing 3.5 μ l aliquots of this solution in each of four Eppendorf tubes labelled A, C, G, and T. Each of these latter four tubes contained 2 μ l of the appropriate Mix solution (i.e. the A tube contained 2 μ l of the A Mix, the C tube contained 2 μ l of the C Mix, etc.). The A, C, G, and T Mix solutions were prepared as described by Messing (1983) with the sole modification that the dideoxy ATP (ddATP) solution used to prepare the A Mix was 0.5 mM in ddATP instead of 2 mM. The polymerase reaction was allowed to proceed for 45 min at 42°C at which point 1 μ l of chase (Messing, 1983) solution was added to each of the four reaction tubes. The incubation was resumed at 42°C for 30 min, followed by heat inactivation of the enzyme at 65°C for 10 min. The samples were evaporated to dryness in a Savant Speed Vac sample concentrator. The pellets were dissolved in 4 μ l of Sequencing Stop Reaction Buffer (100 mM NaOH, 10 mM Na₂EDTA, 0.05% Xylene Cyanol, 0.05% Bromophenol Blue dissolved in 99% de-ionized formamide) and were then either electrophoresed through polyacrylamide gels as described in section II G.2., or stored at -20°C.

(2) T7 DNA polymerase (Sequenase)

DNA sequencing using Sequenase was performed as prescribed by the manufacturer (United States Biochemical Corporation). Briefly, 2.5 μ l of purified "+" DNA (prepared as described in section II.E.2.a.1.b.) was mixed with 2 μ l of the manufacturer's 5X Sequencing Buffer (200 mM Tris·HCl pH 7.5, 100 mM MgCl₂, 250 mM NaCl), 2 μ l of the 0.03 A₂₆₀/ml sequencing primer, and sufficient water to yield a final volume of 10 μ l. The primer was annealed as described in the previous section for Klenow sequencing, after which 1 μ l of 100 mM Dithiothreitol, 2 μ l of a 1:5 dilution of the manufacturer's dGTP based Labelling Mix, 1 μ l of 10 mCi/ml, 600 Ci/mmol [α -³⁵S]dATP, and 2 μ l of a 1:8 dilution of the 13 U/ μ l Sequenase supplied by the manufacturer were added to the primer annealed DNA. The reaction was incubated at room temperature for 5 min before placing 3.5 μ l of it in each of four tubes containing 2.5 μ l of each of the four dGTP based termination mixes supplied by the manufacturer (i.e. one tube contained the dGTP based A termination mix, one the dGTP based C termination mix, etc.). The termination mixes had been pre-warmed at 37°C for 5 min. The termination reactions were incubated at 37°C for 5 min before heat inactivating the enzyme at 65°C for 10 min. The samples were then dried and dissolved in Sequencing Stop Reaction Buffer as described in the previous section for Klenow sequencing.

b) ssDNA sequencing using M13 "+" strand as template

DNA sequencing of purified single stranded M13 "+" strand DNAs was performed directly upon the purified DNA as described in section II.G.1.a.

c) dsDNA sequencing using plasmid as template

DNA sequencing of plasmids involves an initial denaturation step followed by the sequencing reaction (Hefford *et al.*, 1987). The double stranded plasmid DNA molecule was denatured by mixing 1 μ g of plasmid DNA with 1 μ l of 10 mM Na₂EDTA, 4 μ l of 1 M NaOH, and sufficient water to yield a final volume of 20 μ l. The denaturation reaction was incubated at room temperature for 5 min before neutralizing by adding 2 μ l of 2 M NH₄OAc pH 5.4 solution. The denatured DNA was immediately precipitated by adding 55 μ l of cold (-20°C)

95% ethanol to the neutralized denaturation reaction and placing the tube on ice. The DNA was pelleted by microfuging at 13,000 rpm at 4°C for 15 min. The pellet was washed by adding 100 ul of cold (-20°C) 70% ethanol to it, mixing, and microfuging as described above. Again discarded the supernatant and dried the pellet under vacuum in a Savant Speed Vac sample concentrator for 5 min. The pellet was dissolved by incubating it at 37°C for 15 min in a mixture of 2 ul of the desired 0.03 A₂₆₀/ml sequencing primer, 2.0 ul of the 5X Sequencing Buffer supplied by United States Biochemical Corporation for use with Sequenase, and 6 ul of water. This solution was used directly in the labelling reaction in sequencing with Sequenase as described in section II.G.1.a.2.

2. Polyacrylamide gel electrophoresis

Sequencing reactions were analyzed by electrophoresis through denaturing polyacrylamide gels. These gels were either thicker at the bottom (wedge), or of uniform thickness. The gels were electrophoresed at settings of 60w, 60mA, and 1800v on Bethesda Research Laboratories Model S2 DNA sequencers using Sharkstooth combs and 1X TBE Buffer (Biggin *et al.*, 1983). Prior to casting the gels the glass plates were vigorously wiped clean with a paper towel soaked in 95% ethanol. All sequencing reactions were boiled for 2.5 minutes just prior to electrophoresing, and only 2 ul of each sequencing reaction was electrophoresed per loading.

a) Wedge gels

Wedge gels (Carlson and Messing, 1984) were prepared using spacers which were 0.4 mm thick at the top and 1.2 mm thick at the bottom. The gel solution was prepared by mixing 18.75 ml of 40% acrylamide (prepared by dissolving 38 g acrylamide and 2 g of bisacrylamide in 100 ml of water, filtering, and storing the resulting solution in the dark at 4°C), 7.5 ml of 10X TBE Buffer (Biggin *et al.*, 1983), 72 g of electrophoresis grade urea, and sufficient water to yield a final volume of 150 ml. The urea was dissolved by heating the slurry in a hot water bath, and the resulting solution vacuum filtered through Whatman filter paper.

Polymerization was initiated by adding 55 μ l of TEMED, 550 μ l of 10% Ammonium Persulphate and swirling. The gel was then cast and left to polymerize for 1 h.

The polymerized gel was mounted in the apparatus and pre-electrophoresed for 30 min prior to loading the samples. The gel was run until the Bromophenol Blue dye was just short of the bottom of the gel. The gel sandwich was removed from the sequencing apparatus and the gel floated in 5% methanol, 5% acetic acid solution for 30 min prior to transferring it to Whatman 3MM paper and drying. The dried gel was exposed to Kodak X-Omat RP X-ray film for 2 to 4 days before developing the film.

b) Uniform thickness gels

Uniform thickness gels (Carlson and Messing, 1984) were prepared using 0.4 mm spacers. The gel solution was prepared by mixing 15 ml of 40% acrylamide (see the previous section), 10 ml of 10X TBE Buffer (Biggin *et al.*, 1983), 48.1 g of solid electrophoresis grade urea, and sufficient water to yield a final volume of 100 ml. The solution was heated and filtered as described for wedge gels. Polymerization was initiated by adding 40 μ l of TEMED and 400 μ l of 10% Ammonium Persulphate and swirling. The gel was then cast and left to polymerized for 1 h.

The polymerized gel was mounted in the apparatus and pre-electrophoresed for 30 min. After sample loading the gel was electrophoresed until the desired number of nucleotides had been run off the bottom of the gel. This was estimated by following the migration of the marker dyes. The gel sandwich was removed from the sequenator apparatus and the gel soaked in 5% methanol, 5% acetic acid solution for 10 min prior to transferring it to Whatman 3MM paper and drying. The dried gel was exposed as described in the previous section for wedge gels.

c) Gel autoradiography

The dried sequencing gels were placed in X-ray cassettes along with undeveloped Kodak X-Omat RP X-ray film. The film was placed in direct contact with the dried gel, and left to

expose for 2 to 4 days at room temperature with lead weights of top of the cassette. The film was developed as recommended by the manufacturer.

3. Sequence analysis

DNA sequence was analyzed using the Pustell Sequence Analysis Program supplied by International Biotechnologies Incorporated.

II. Heterologous expression in *E. coli*

1. Construction of expression clones

All expression clones, whether based on lambda gt11, pUC 19, or pGEM-2 vectors, were constructed and had their DNA sequence verified using the protocols for DNA isolation and manipulation described in section II.E and II.G.

2. Sample preparation and analysis

a) Protein sample preparation

(1) pUC 19 based expression clones

Protein samples from cultures of *E. coli* harboring expression plasmids were prepared in several ways. The simplest method involved growing an overnight culture in 10 ml of rich medium (i.e. either L or 2X YT Medium, see Appendix A) containing a 1X concentration of the appropriate antibiotic and being 2 mM in IPTG (the *lac* operon inducer). After growth, the cells were pelleted and washed twice in 1 ml of *E. coli* Lysis Buffer (50 mM Tris·Cl pH 8, 1 mM Na₂EDTA, 100 mM NaCl). The final pellet was resuspended in 200 ul of *E. coli* Lysis Buffer and the cells lysed by adding an equal volume of 2X Laemmli Buffer (100 mM Tris·Cl pH 6.8, 4% SDS, 10% β-mercaptoethanol, 20% glycerol, 0.02% Bromophenol Blue) and boiling for 5 min. The preparation was stored at 4°C until needed.

More sophisticated methods involved inoculating 25 ml of rich medium (either L or 2X YT Medium) in a 250 ml Erlenmeyer flask containing a 1X concentration of the appropriate antibiotic with 250 ul of an overnight culture of *E. coli* harbouring the expression clone of interest. This culture was incubated with agitation at 225 rpm at 37°C in a New Brunswick Scientific Model G24 Environmental Shaker Incubator for 3.5 h (i.e. mid-log phase) at which

point 500 μ l of 100 mM IPTG (Appendix A) was added to the culture and the incubation continued for 2 h. The cells were then harvested at 4000 rpm for 5 min at 4°C in the Beckman JA-20 rotor, and the cell pellet washed once with 3 ml and once with 1 ml of *E. coli* Lysis Buffer before snap freezing the final pellet in liquid nitrogen. The frozen pellet was thawed while vortexing at the maximum setting on a Fisher Scientific Vortex-Genie in 500 μ l of *E. coli* Lysis Buffer containing 1.3 μ l of a 50 mM Phenylmethylsulphonyl Fluoride (PMSF, Sigma) stock solution (dissolved in methanol). The lysis procedure was completed by adding 500 μ l of 2X Laemmli Buffer and boiling for 5 min. A 20 μ l aliquot of this final solution was usually subjected to SDS-PAGE analysis as described in section II.H.2.b.

The most detailed protein sample preparation involved a fractionation of transformed *E. coli* cultures into extracellular, intracellular soluble, and intracellular insoluble protein fractions. Cultures (25 ml) were grown and induced as described above. The cells were pelleted at 8000 rpm for 10 min at 4°C in the Beckman JA-20 rotor, and the cell pellet and supernatant separated and saved.

The supernatant was recentrifuged as described above, and the pellet discarded. A 20 ml aliquot of the resulting supernatant was cooled to 0°C and the proteins precipitated by mixing in 5 ml of cold (4°C) 100% trichloroacetic acid solution by vortexing at a setting of 3, followed by incubation on ice for 45 min. The precipitate was collected by centrifuging at 8000 rpm for 15 min at 4°C in a Beckman JA-20 rotor. After discarding the supernatant, the pellet was air dried overnight before dissolving it in 100 μ l of 1X Laemmli Buffer and boiling for 5 min.

The cell pellet fraction was resuspended in 1 ml *E. coli* Lysis Buffer and repelleted by centrifuging at 5000 rpm in a Beckman JA-20 rotor for 10 min at 4°C. The cells were lysed by adding glass beads, 1 ml of *E. coli* Lysis Buffer containing 5 mg/ml lysozyme, and 2.6 μ l of 50 mM PMSF stock solution to the pellet and passing this mixture through three freeze-thaw cycles while vortexing at the maximum setting. The resulting frothy solution was centrifuged at 8000 rpm for 15 min at 4°C in the Beckman JA-20 rotor. After separation of the pellet

and the supernatant an equal volume of 2X Laemmli Buffer was added to the intracellular soluble protein containing supernatant fraction, and the solution was boiled for 5 min. The pellet, which contained the intracellular, insoluble proteins, was washed by resuspending in a mixture of 5 ml of *E. coli* Lysis Buffer and 5 ml of TE Buffer. The proteins were pelleted as before, and then dissolved in 0.5 ml of 1X Laemmli Buffer by vortexing at a setting of 5, boiling for 5 min, and repeating the vortex and centrifugation steps. The resulting supernatant fraction was considered as containing the intracellular, insoluble proteins.

(2) pGEM-2 based expression clones

(a) Preparation of unlabelled proteins

The pGEM-2 based expression clones are constructed such that the TIV sequences are fused behind the Bacteriophage T7 RNA Polymerase promoter sequence encoded by this plasmid. Heterologous expression of these sequences in *E. coli* requires transformation of the recombinant plasmid into a cell line which already contains a plasmid encoding the polymerase as described by Tabor and Richardson (1985). *E. coli* strain K38 containing plasmid pGP1-2 was a generous gift of Dr. S. Tabor (Harvard University). The cells were transformed with the recombinant pGEM-2 based plasmid as described in section II.D. Transformed cells were selected by their ability to grow on petri dishes containing 50 ug/ml of both Ampicillin and Kanamycin. The presence of both plasmids within the cells in these colonies was verified by restriction endonuclease analysis (see section II.E.1.e.) of plasmid DNA mini preps (see section II.E.1.d.) of the selected colonies.

The production of unlabelled proteins involved inoculating 10 ml of T7 Expression Medium (2% Bactotryptone, 1% Yeast Extract, 0.5% NaCl, 0.2 % Glycerol, 50 mM KPO_4 pH 7.2) containing 50 ug/ml of both Ampicillin and Kanamycin with 200 ul of an overnight culture of *E. coli* K38 bacteria containing both plasmids. The ensuing culture was incubated at 30°C with agitation at 180 rpm in a New Brunswick Scientific Controlled Environment Incubator Shaker until the A_{590} reached 1.5. The temperature was then shifted to 42°C and the incubation continued for 25 min. Host cell RNA Polymerase activity was then inhibited

by adding 20 mg/ml Rifampicin (Sigma) stock solution dissolved in methanol to a final concentration of 100 ug/ml. After incubating the culture for 2 h at 37°C with agitation at 200 rpm in a New Brunswick Scientific Model G24 Environmental Incubator Shaker, the cells were collected by centrifugation at 5000 rpm for 5 min in a Beckman JA-20 rotor at 4°C. The cell pellet was washed by resuspending it in 1 ml *E. coli* Lysis Buffer, and then repelleting the cells. The final cell pellet was resuspended in 200 ul of *E. coli* Lysis Buffer, and the proteins liberated by the addition an equal volume of 2X Laemmli Buffer and boiling for 5 min. Both 15 and 30 ul aliquots of the resulting samples were subjected to SDS-PAGE analysis as described in section II.H.2.b.

(b) Preparation of ³⁵S labelled proteins

The preparation of ³⁵S labelled proteins using this two plasmid system was performed as described by Tabor and Richarson (1985). A 100 ul aliquot of an overnight culture of *E. coli* strain K38 harbouring both pGP1-2 and the pGEM-2 based plasmid was used to inoculate 10 ml of L Medium (Appendix A) containing 40 ug/ml of both Ampicillin and Kanamycin. The resulting culture was incubated at 30°C with agitation at 180 rpm in a New Brunswick Scientific Controlled Environment Incubator Shaker until the A₅₉₀ reached 0.5. At this point 0.2 ml of the culture was transferred to a 13x100mm culture tube and the cells pelleted by centrifugation at 5000 rpm in a Beckman JA-20 rotor for 5 min at 4°C. The cells were washed by suspending in 5 ml M9 media (Maniatis *et al.*, 1982), and then pelleted as before. The cells were resuspended in 1 ml of M9 media supplemented with 20 ug/ml thiamine and 0.6 μl 18 amino acids (minus cysteine and methionine) and incubated at 30°C with agitation at 180 rpm for 60 min. The temperature was then shifted to 42°C and the incubation continued for 15 min. Host cell RNA polymerase activity was inhibited by the addition of 10 ul of 20 mg/ml Rifampicin stock solution. The incubation at 42°C was continued for 20 min prior to shifting the temperature to 30°C for 20 min. The proteins were labelled by the addition of 15 uCi of ³⁵S L-methionine (15 uCi/ul, 800 Ci/mmol, Amersham) and incubating at 30°C for 5 min. The culture was then transferred to an Eppendorf tube and the cells

pelleted by microfuging at 13,000 rpm for 20 seconds at room temperature. The cell pellet was resuspended in 120 μ l of cracking buffer (60 mM Tris·HCl pH 6.8, 1% SDS, 1% 2-mercaptoethanol, 10% glycerol, and 0.01% Bromophenol Blue) by vortexing at a setting of 5 on a Fisher Scientific Vortex-Genie. The samples were heated at 95°C for 3 min and 20-60 μ l aliquots subjected to SDS-PAGE analysis as described in section II.I.3.

b) SDS-PAGE analysis of samples

Polyacrylamide gel electrophoresis of protein samples was performed as described by Laemmli (1970). The gels were cast using 16cm x 18cm glass plates separated by a 1.5mm spacer with a 10% acrylamide base and a 4% acrylamide stack prepared as described in Appendix B. Both gel solutions were degassed for 15 min under vacuum prior to initiating polymerization with TEMED and Ammonium Persulphate. Electrophoresis was performed at 45 v for 16 h using 25 mM Tris·Cl pH 8.3, 0.1% SDS, 192 mM glycine solution as running buffer. The protein bands were visualized by staining the gel in 0.1% Coomassie Brilliant Blue R in 40% methanol, 10% acetic acid solution for 1 h, followed by destaining in 40% methanol, 10% acetic acid solution for 5 to 20 h. Destained gels were photographed through an orange filter under white light using Polaroid Type 52 Polapan Film in a Polaroid Land Camera.

c) Western blotting

SDS-PAGE gels to be Western blotted (Tobin *et al.*, 1979) were electrophoresed as described in the previous section except that no staining was done. The electrophoresed samples were transferred immediately to nitrocellulose (Schleicher & Schuell) using the Hoefer Model TE50 apparatus and 25 mM Tris·Cl pH 8.3, 192 mM glycine, 20% v/v methanol solution as transfer buffer. The transfer was performed at 4°C for one hour at 32.5v/250mA, followed by a second hour at 55v/400mA. The blotted filter was rinsed briefly (2 minutes) in TBST (10 mM Tris·Cl pH 8, 150 mM NaCl, 0.05% Tween 20) and the protein bands detected using the Promega BioTec ProtoBlot Immunoblotting System. All incubations involved in the immunodetection of the blotted proteins were performed at room

temperature. In order to block any unblocked sites on the membrane prior to probing with the antibody the washed filter was agitated in TBST containing 1% Bovine Serum Albumin for 30 min. This was followed by removal of the blocking solution and binding of the primary antibody by agitating the filter for 30 min in 50 ml TBST containing the appropriate dilution of the rabbit anti- β -galactoside antibody. This solution was then poured off, the filter rinsed briefly (10 sec) in TBST before washing for 5, 10, and 5 min with agitation in TBST. The washed filter was incubated for 30 min with agitation in 50 ml TBST containing an appropriate dilution of the secondary antibody (Anti-Rabbit IgG (Fc) Alkaline Phosphatase Conjugate) which had been incubated with an *E. coli* lysate (Helfman *et al.*, 1983). The *E. coli* lysate was prepared by growing a 150 ml overnight culture of the *E. coli* strain of interest. After centrifugation, the pelleted cells were resuspended in 10 ml TBST, boiled for 15 min, and the resulting lysate was divided into 1 ml units for immediate storage at -70°C . The secondary antibody was precompeted with this lysate by mixing 400 μl of the thawed lysate with an equal volume of TBST, and 6.7 μl of the Anti-Rabbit IgG (Fc) Alkaline Phosphatase Conjugate. The mixture was rocked gently for 2 h at 4°C followed by 1 h at room temperature before clarifying the suspension by microfuging for 5 min at room temperature. The resulting supernatant was used for specific binding of the secondary antibody to the rabbit antibodies bound to the filter as described above. Upon completion of the binding step, the filter was washed three times in TBST for 5, 10, and 5 min, respectively. The washed filters were "developed" by placing them in 30 ml of Alkaline Phosphatase Buffer (100 mM Tris $\cdot$ Cl pH 9.5, 100 mM NaCl, 5 mM MgCl_2) containing 198 μl of 50 mg/ml Nitro Blue Tetrazolium in 70% dimethylformamide stock solution, and 99 μl of 50 mg/ml 5-bromo-4-chloro-3-indolyl phosphate in dimethylformamide stock solution, and gently hand mixing. When the blue-purple colour of the positive bands reached sufficient intensity that the bands could be easily visualized the reaction was quenched by placing the filter in water.

I. *In vitro* synthesis of RNA and protein

1. Synthesis of RNA *in vitro*

a) *In vitro* transcription

The synthesis of RNA *in vitro* was performed using the protocol recommended by the manufacturer of the Bacteriophage T7 RNA Polymerase (Promega Biotec). The synthesis step involved the addition of 1 ug of RNase-free dsDNA template (containing the Bacteriophage T7 RNA Polymerase promoter) to a solution containing 2 ul of 10X Bacteriophage T7 RNA Polymerase Buffer (400 mM Tris·Cl pH7.9, 60 mM MgCl₂, 100 mM NaCl, and 20 mM spermidine), 2 ul of 100 mM DTT, 1 ul of a 10 mM solution of each of the four NTPs (prepared as described by Maniatis *et al.*, 1982), sufficient water to yield a final volume of 20 ul, and 0.5 ul of 15 U/ul Bacteriophage T7 RNA Polymerase. The reaction was incubated at 37°C for 30 min, at which point the enzyme was inactivated by heating at 65°C for 5 min. The resulting RNA solution was stored at -70°C until needed.

When the *in vitro* synthesized RNA was to be used for *in vitro* translation, a second 0.5 ul aliquot of the enzyme was added to the reaction mixture after the initial 30 min reaction period and the reaction allowed to continue for a further 30 min at 37°C. The enzyme was then inactivated by heating as described above. The RNA was purified by extracting a pool of three reactions for one DNA template once with one volume phenol/chloroform (see section II.E.1.a.). After removal of a 4.5 ul aliquot, which was used for agarose gel analysis and estimation of concentration, a 1/20th volume of 2.5 M NaOAc pH 5.2 solution was added to the separated aqueous phase. After mixing by gentle hand inversion, the RNA was precipitated by adding 2.5 volumes of cold (-20°C) ethanol, mixing by gentle inversion, and incubating overnight at -20°C. The precipitate was collected by microfuging at 13,000 rpm at 4°C for 30 min. The ensuing pellet was washed with 70% ethanol and dried as described in section II.E.1.a. before dissolving it in 12 ul of water. The RNA concentration of the resulting solution was estimated from agarose gel analysis of the aliquot removed prior to precipitation of the RNA, and the solution was stored at -70°C until needed.

b) Agarose gel analysis of RNA

Agarose gel analysis of RNA samples were performed using agarose-formaldehyde gels as described by Maniatis *et al.* (1982). The desired amount of solid agarose (1.0 to 1.5% final) was dissolved in the appropriate volume of water by heating. After allowing the ensuing solution to cool to 60°C, the appropriate volume of 10X MOPS Buffer (0.2 M MOPS (3-[N-Morpholino]-propanesulfonic acid) pH 7.0, 50 mM sodium acetate, and 10 mM Na₂EDTA pH 8.0), and pure formaldehyde (to 2.2 M final concentration) were added and the gel cast. Samples were prepared by adding 4.5 ul of the RNA solution of interest to 2 ul of 10X MOPS Buffer, 3.5 ul of formaldehyde, 10 ul of de-ionized formamide, and 2 ul of 10X RNA Sample Loading Buffer (50% glycerol, 1 mM Na₂EDTA pH 8.0, 0.4% Bromophenol Blue, and 0.4 % Xylene Cyanol). The samples were then heated at 65°C for 5 min and electrophoresed. After electrophoresis the bands were visualized by staining for 20 min in 1 ug/ml EtBr solution, destaining for 20 min (or longer if necessary) in water, and photographing under short wavelength (254 nm) ultraviolet light through a red filter.

2. *In vitro* translation

In vitro translations were performed utilizing the Rabbit Reticulocyte Lysate *in vitro* protein synthesizing system as recommended by the manufacturer (Amersham). The synthetic step was carried out by adding 0.25-0.75 ug of RNA produced *in vitro* (see section II.I.1.a.) to 20 uCi ³⁵S labelled L-methionine (15 uCi/ul, 800 Ci/mmol, Amersham), 16 ul of the Rabbit Reticulocyte Lysate, and sufficient water to yield a final volume of 20 ul. The reaction was incubated at 30°C for 1 h, after which an equal volume of 2X Laemmli Buffer (see section II.H.2.a.1.) was added to each sample. The samples were incubated in a boiling water bath for 5 min prior to electrophoresis on SDS-PAGE gels as described in the following section.

3. SDS-PAGE analysis of ³⁵S labelled proteins

SDS-PAGE gels for the analysis of ³⁵S labelled proteins (prepared in sections II.H.2.a.2.b and II.I.2.) were prepared and electrophoresed exactly as described in section II.H.2.b. The

gels were fixed in 30% methanol, 10% acetic acid solution for 30 min, and then incubated in Amplify (Amersham) for 30 min prior to their transfer to 3MM Whatman filter paper and drying under vacuum at 60°C. The dried gels were exposed to Kodak XAR film at -70°C for 1 to 4 days prior to developing the film as recommended by the manufacturer.

III. RESULTS

A. History of the L10 plasmid carrying the L96 gene

As described in section I.F the sequencing of a second TIV gene and the comparison of the coding sequence and possible consensus regulatory sequences of this gene with those of the recently described TIV capsid gene (Tajbakhsh, 1988) should provide much needed information on the mechanisms involved in TIV late gene transcription. Detection of sequence within the L96 gene was fortuitously made by Tajbakhsh (1988) through use of a synthetic DNA hybridization probe to clone the major TIV capsid gene. In these experiments the oligonucleotide probe sequence was deduced by back translation of a peptide fragment of the major TIV capsid protein. Due to the degeneracy of the genetic code of this particular peptide, the synthetic oligomer was made as a mixture of 192 possible combinations (Figure 1A). As shown in Figure 1B to 1D the 20-mer mixed oligonucleotide probe, when labelled with [γ - 32 P]ATP and hybridized with TIV DNA digested with different restriction endonucleases (Figure 1B), hybridizes to several DNA fragments depending on whether the post hybridization washes were at low (Figure 1C) or high (Figure 1D) stringency. Further screening of a partial TIV DNA bank constructed in the Bam HI site of plasmid pBR322 with this probe revealed two positive plasmids: L1 (15.8 kbp) and L10 (16 kbp) containing BamHI fragments 11.5 and 11.8 kbp in size, respectively (Figure 1E). However, additional experiments by Tajbakhsh (1988) with this oligonucleotide mixture at high stringency washes showed that neither L10 nor L1 TIV DNA carried capsid-specific sequence.

Subsequent experiments localized the hybridization of the mixed oligonucleotide probe to a 1.05 kbp SalI/HindIII fragment of L10, while the hybridization to L1 proved to be an artifact. Further use of the mixed oligonucleotide probe as a sequencing primer for the 1.05 kbp SalI/HindIII band indicated that there was no match with the original capsid protein peptide fragment used in the back translation. However, preliminary evidence that the 1.05 kbp SalI/HindIII fragment has coding function was revealed when this fragment was

Figure 1. Cloning of the L96 gene from the TIV genome by oligonucleotide hybridization.

(A) Amino acid sequence of peptide fragment 11-4 of the major TIV capsid protein, and the corresponding back translated DNA sequence (oligonucleotide 11-4). TIV DNA was digested with various restriction endonucleases, electrophoresed in 0.7% agarose (B) and alkaline transferred to a nylon membrane. Oligonucleotide 11-4, end labelled with ^{32}P , was hybridized to this blot and the blot was washed under low (C) and high (D) stringency conditions (RT and 50°C , respectively). BamHI fragments #10 (10.8 kbp) and #6 (14.3 kbp), EcoRI fragments #10 (8.5 kbp) and #1 (approx. 25 kbp), and Sall fragments #33 (2.0 kbp), #23 (4.8 kbp), and #11 (10.2 kbp) hybridized to the probe with low stringency washes only. BamHI #10 (10.8 kbp), EcoRI #1 (approx. 25 kbp), and Sall #33 (2.0 kbp) hybridized after the high stringency wash (see Tajbakhsh, 1986 and 1988 for a description of the restriction endonuclease banding pattern of TIV DNA). (E) Hybridization of the oligonucleotide to a dot blot of pBR322 clones containing BamHI fragments of TIV DNA (low stringency wash). (F) Uninfected (12 h p.i.) and infected *Estigmene acrea* total RNA (6 ug) were electrophoresed in a 1.0% agarose-formaldehyde gel, Northern blotted to nylon membrane in 20X SSC solution, and hybridized with a nick translated 1.05 kbp Sall/HindIII fragment clone from L10 (Tajbakhsh, 1988).

M = HindIII digested lambda DNA molecular weight markers

B = BamHI

E = EcoRI

S = Sall

A

Tyr Met Tyr Gly Gly Tyr Thr Ala Thr Ala Tyr Thr

↓

TAT ATG TAT GGT GGT TGT A

C C C C C
A A G
G G

E

F

found to hybridize specifically to a 3.6 kb transcript in total RNA obtained from *Estigmene acrea* cells 8 to 24 h after infection with TIV (Figure 1F). This putative gene contained on clone L10 was called LW1 was set aside at that time for future characterization. Since the sequences carried by L10 are for a late gene (Figure 1F), comparison of its potential regulatory sequences with those of the late expressing capsid gene sequenced by Tajbakhsh (1988) should help in the elucidation of possible TIV late gene consensus regulatory sequences. During the course of this work the LW1 gene located on clone L10 was given the more descriptive name L96 based on the putative size of the protein predicted from the DNA sequence reported in section III.C.

B. Physical mapping of the L96 gene

1. Restriction site map of clone L10

The first requirement for characterization of the L96 gene was the preparation of an accurate restriction site map of clone L10. This was obtained by digesting purified L10 DNA with the appropriate restriction endonucleases--either singly or in pairs--and then electrophoresing the resulting DNA fragments through an agarose gel (Figure 2A). Analysis of the resulting banding pattern, in conjunction with the known restriction sites in the pBR322 region of clone L10, were used to deduce the restriction endonuclease map for L10 shown in Figure 2B. Although clone L10 was constructed by ligating BamHI^I digested TIV DNA into the BamHI site of pBR322, the insert in this clone could not be excised by BamHI digestion. Specifically the right-hand BamHI site is not recognized by the enzyme, presumably due to a deletion which occurred in the cloning process. This theory is supported by the observation that fragments of the TIV insert in L10 exclusively hybridize to a BamHI fragment larger than 11.8 kbp (approximately 18.65 kbp, see next section).

2. Mapping of the L96 gene on the L component of the TIV genome

Another goal of this thesis is to increase the understanding of the organization of the TIV genome. The DNA fragments from restriction endonuclease digestion of total TIV DNA with seven different enzymes have been determined (Tajbakhsh *et al.*, 1986; Tajbakhsh,

Figure 2. Restriction endonuclease mapping of clone L10.

(A) L10 DNA was digested with various restriction endonucleases and electrophoresed in 0.7% agarose. (B) Restriction endonuclease map of L10 deduced from analysis of the gel in (A). Restriction enzymes which were observed not to cleave L10 DNA include XbaI and XhoI, while BamHI and PstI were found to have no sites located within the TIV insert in L10.

M = HindIII digested lambda DNA molecular weight markers

B = BamHI

B⁰ = null BamHI site

Bg = BglII

E = EcoRI

H = HindIII

Hp = HpaI

P = PstI

S = SalI

X = XbaI

Xh = XhoI

[illegible]

B

L 10

1988); hence, via the hybridization of selected probes derived from the TIV insert in L10 to replica blots of restriction endonuclease digested total TIV DNA, an extended restriction site map of the TIV genome surrounding the L96 gene may be obtained. The resulting map will eventually be used to obtain the restriction site map of the complete L component of the TIV genome (an ongoing objective in this laboratory). The experiment was performed in duplicate with one blot being probed with a fragment from one end of the insert in L10, and the second blot being probed with a fragment from the opposite end of the L10 insert. In order to facilitate analysis of the autoradiograms the probes were chosen such that, as much as was possible, only one fragment was selected in each digest--namely the fragment which spans the end of the TIV insert in L10. Via analysis of the restriction endonuclease map of L10 shown in Figure 2B, the 5.5 kbp BamHI/SalI fragment from the 3' end of the insert in L10 and the 1.8 kbp SalI fragment of L10 which ends at the SalI site contained within the vector derived region of L10 were selected as the 3' and 5' probes to map the restriction endonuclease sites 3' and 5' of the L10 insert, respectively (Figure 3A). The 3' and 5' designations refer to the orientation of the L96 gene.

The 5.5 kbp BamHI/SalI fragment had previously been cloned into the BamHI/SalI sites of plasmid pUC 18. Since Tajbakhsh *et al.* (1986) had shown that no homology existed between pBR322 and TIV sequences, and pUC 18 is a derivative of pBR322, the entire recombinant plasmid was used as probe for the 3' end. The 1.8 kbp SalI fragment of L10 was purified from SalI digested L10 DNA by electroelution. Both probes were labelled with [α - 32 P]dATP by nick translation and used to probe identical replica blots of restriction endonuclease digested total TIV DNA (Figure 3B). The information contained in the resulting autoradiograms (Figure 3C and D) is summarized in Table IV, and was used to construct the extended restriction endonuclease map shown in Figure 3E.

Both probes hybridized to single bands of identical sizes in the BamHI, PstI, XbaI, and XhoI digests. Clearly these bands, which are 18.65, >23.1, 20.75, and 12.75 kbp in size, respectively, are the genomic fragments on which the TIV insert in L10 is located. The fact

Figure 3. Mapping of the L96 gene on the TIV genome.

(A) The 5.5 kbp BamHI/Sall and 1.8 kbp Sall fragments were selected for use as probes to determine the restriction sites adjacent to the left and right BamHI sites, respectively, of the L10 TIV insert. The dotted line at the right end of the 1.8 kbp Sall fragment denotes that this fragment terminates at the Sall site in the pBR322 region of L10. TIV DNA was digested with restriction endonucleases, electrophoresed in 0.7% agarose (B), alkaline transferred to nylon membrane, and probed with either nick translated 5.5 kbp BamHI/Sall fragment (C) or nick translated 1.8 kbp Sall fragment (D). (E) Extended restriction endonuclease map of the TIV genome in the region of the L96 gene deduced from analysis of the autoradiograms in (C) and (D). For sites outside the TIV insert in L10, only those restriction fragments containing a corresponding site within the boundaries of the insert in L10 are indicated. Note that the location of the left-most BglII, EcoRI, and HpaI sites could not be placed on this map for the reasons explained in the text.

M = HindIII digested lambda DNA molecular weight markers

B = BamHI

B⁰ = null BamHI site

Bg = BglII

E = EcoRI

H = HindIII

Hp = HpaI

P = PstI

S = Sall

X = XbaI

Xh = XhoI

L10

Table IV. Extended restriction site mapping of the TIV genome in the region of the L96 gene: fragments selected by hybridization to 3' and 5' probes derived from L10, and deduced distances from the termini of the TIV insert in L10.^a

restriction endonuclease	3' probe (size kbp)	kbp from 3' end of insert	5' probe (size kbp)	kbp from 5' end of insert
BamHI	18.7	b	18.6	b
BglII	8.5	c	8.6	0.1
	5.3	?		
	4.8	?		
EcoRI	8.7	d	8.8	d
	3.1	?	1.75	1.15
	2.55	?		
HindIII	6.0	d	2.8	1.8
	5.4	5.0	1.85	d
HpaI	6.9	?	6.0	1.1
	6.4	?		
PstI	>23.1	b	>23.1	b
Sall	15.7	10.2	10.5	9.2
XbaI	21.0	b	20.5	b
XhoI	12.5	b	13.0	b
Sall/XhoI	-	-	1.5	0.16

^a The 3' and 5' designation are relative to the orientation of the L96 gene.

^b Indicates that the selected fragment contains the entire TIV insert in L10.

^c Indicates that the selected fragment is identical to that observed with the 5' probe.

^d Indicates that the selected fragment is contained within the TIV insert in L10.

that the genomic BamHI fragment is larger than the TIV insert in L10 provides one possible explanation for the null BamHI site found in L10. Specifically, a portion of the genomic BamHI fragment containing the L96 gene was deleted in the cloning process, and that the cloning of the insert in L10 into the BamHI site of pBR322 was fortuitous. Since PstI, XbaI, and XhoI do not cut within the L10 insert (Figure 2), these results indicate that the single fragment selected in each of these digests contains the entire L10 insert.

The autoradiograms were analyzed by comparison of the sizes of the selected bands to the known restriction endonuclease map of L10. For example, the only SalI fragment the 3' probe would be predicted to hybridize to is the one spanning the 3' BamHI site. Experimentally a single fragment approximately 15.7 kbp in size is observed to hybridize to the probe; hence, the adjacent SalI site is 10.2 kbp to the left of the BamHI site. Similarly in the HindIII digest two bands would be expected to hybridize--the 6.2 kbp fragment contained completely within the L10 insert, and the HindIII fragment of unknown size spanning the 3' BamHI site of the insert. Experimentally two bands of size 5.4 and 6.0 kbp are observed to hybridize to the probe, hence the adjacent HindIII site is located approximately 5 kbp to the left of the BamHI site. Like the HindIII digest, the EcoRI digest would be predicted to yield two positive bands--one 8.8 kbp in size which is completely contained within the L10 insert and one of unknown size spanning the 3' BamHI site. However, experimentally three bands are observed to repeatedly hybridize to the probe. The most strongly hybridizing band was approximately 8.7 kbp in size and is the aforementioned 8.8 kbp fragment. The remaining positive fragments were of equal, but lesser, intensity and were approximately 2.55 and 3.1 kbp in size. Similar problems were encountered in the BglII and HpaI digests with the 3' probe. As a consequence the location of these sites 3' of the 3' BamHI site could not be placed on the extended restriction endonuclease map. These problems are difficult to explain--one possible cause is incomplete digestion of the DNA by these enzymes. However, this is unlikely as the results described above are repeatable. The results obtained with the 5' probe were analyzed in the same way as described above for the 3' probe, and the data from

both analyses were used to construct the extended restriction endonuclease map shown in Figure 3E.

In order to try and place the PstI, XbaI, and XhoI sites on the extended restriction endonuclease map, and to investigate the null BamHI site in L10, double digests of these four enzymes with Sall were included in the experiment. Unfortunately, with the exception of the probing of the Sall/XhoI codigest with the 5' probe, all of these digests yielded a greater number of positive bands than expected, and therefore made analysis impossible. The 1.5 kbp fragment selected in the Sall/XhoI codigest with the 5' probe places the XhoI site 0.16 kbp 5' of the 5' end of the insert in L10. Hence, since the genomic XhoI fragment on which the L10 insert is located is 12.75 kbp long, the 3' XhoI site is located 1.8 kbp 3' of the 3' BamHI site. Although the BamHI/Sall codigest was predicted to produce only one band which would hybridize to the 5' probe, experimentally three bands (16.0, 9.6, and 7.9 kbp) were observed to hybridize (Figure 3D). All of these bands are greater than 7.9 kbp in size thereby placing the 5' BamHI site at least 6.6 kbp 5' of the 5' end of the TIV insert in L10. This provides further evidence in support of the theory that the null BamHI site in L10 arose due to a deletion during the cloning process.

C. Analysis of the L96 gene

1. Preparation of deletion clones for DNA sequencing

As described in section III.A. preliminary sequence data showed the 1.05 kbp HindIII/Sall fragment of L10 (see Figure 4) to possess an open reading frame (ORF) running the length of the fragment from the HindIII site to the Sall site (Tajbakhsh, 1988). No ORFs running the length of this fragment in the Sall to HindIII direction were found; hence, the orientation of the gene was from the HindIII to Sall sites. Moreover, when this fragment was used to probe a Northern blot of total RNA from TIV-infected *Estigmene acrea* cells a 3.6 kb transcript was exclusively identified (Tajbakhsh, 1988). Since the location of the 1.05 kbp HindIII/Sall fragment on the 3.6 kb transcript was not known, it was decided to sequence outwards from both the HindIII and Sall ends looking for suitable initiation and termination

Figure 4. Subcloning of L10 restriction endonuclease fragments into M13mp19 for DNA sequencing purposes.

The fragments 3' and 5' of the 1.05 kbp fragment of the L96 ORF were isolated by the appropriate restriction endonuclease digestion followed by electroelution of the desired band. The 3.0 kbp EcoRI/Sall fragment 3' of the ORF was cloned directly into the EcoRI/Sall sites of M13mp19 to yield clone W3. The TIV insert from this clone was then isolated with blunt/Sall termini by digestion of W3 RF with EcoRI, followed by Klenow treatment to fill in the recessed 3' termini, Sall digestion and electroelution of the 3.0 kbp band. Ligation of this fragment to M13mp19 which had been digested with HindIII, Klenow treated, and digested with Sall yielded clone W4 with the insert in the opposite orientation relative to W3. The 2.5 kbp HindIII fragment 5' of the ORF was isolated by HindIII digestion of L10 followed by electroelution of this band. Ligation to HindIII digested M13mp19 RF produced a population of clones with the insert in both orientations. Clones W5 and W6 were differentiated on the basis of restriction endonuclease analysis.

B = BamHI

B⁰ = null BamHI site

Bg = BglII

E = EcoRI

(E) = EcoRI site not regenerated in the subcloning

H = HindIII

Hp = HpaI

S = Sall

L10

Figure 5. Sequencing strategy for the L96 gene.

(A) Sizing of M13mp19 Dale deletion clones. Phage containing supernatants were prepared, the phage lysed, and the released DNA electrophoresed in a 0.7% agarose gel as described in materials and methods. Full size (W4) and vector (M13) clones were also run as markers.

(B) Sequencing strategy of both strands of the TIV genome in the region of the L96 gene using Dale deletion clones and subcloned fragments.

A

B

codons. From the size of the transcript the maximum possible amount required at either end was 2.6 kbp; hence, the 3 kbp EcoRI/SalI fragment at the 3' end of the ORF and the 2.5 kbp HindIII fragment at the 5' end of the ORF were selected for subcloning (Figure 4). Both fragments were cloned in both orientations into the single-stranded vector M13mp19. Overlapping deletion subclones (Figure 5) of all four of these parental subclones were prepared using the method of Dale *et al.* (1985). These clones, along with several subclones of L10 provided by Dr. S. Tajbakhsh, were used for DNA sequencing.

2. DNA sequence

a) The L96 ORF

The deletion clones and subclones of L10 prepared in the previous section were sequenced using the dideoxy chain termination method of Sanger *et al.* (1977). This yielded a block of 4254 bp of the TIV insert in clone L10 which was sequenced in both directions (Figure 6A). Analysis of this sequence information revealed a single major open reading frame. This ORF, along with any others encoding a putative protein greater than 60 amino acids in length in the region sequenced, were localized on the TIV insert in clone L10 (Figure 7). From this analysis clearly the 2604 bp L96 ORF is the major one in this region of the TIV genome, although ORFs D, E, H, I, and K encode putative proteins of significant size.

b) Sequence characteristics

The sequenced region was found to contain 36.8% adenine, 17.7% cytosine, 16.3% guanine, and 29.2 % thymine residues. These results agree with those of Tajbakhsh (1988) who reported the TIV capsid gene to be A-T rich. Another interesting feature of the sequenced region was the presence of numerous (>50) direct repeats of at least 12 nucleotides in length. The longest of these sequences was 46 nucleotides in length. Interestingly, approximately 80% of these direct repeats are concentrated in a region 1350 to 1800 bp downstream of the putative L96 start codon. One of these direct repeats, 22 nucleotides in length, was repeated three times within this region. The location of this cluster

Figure 6. Nucleotide sequence and deduced amino acid sequence of the TIV L96 gene and gene product.

(A) L96 gene and amino acid sequences. (B) Hydropathy plot of the L96 gene product using the parameters specified by Kyte and Doolittle (1982) with a window size of 9.

AATAAA = eukaryotic polyadenylation signal

CAT = putative CAT box

IR = inverted repeats

TATA = putative TATA box

A

60

TTTCGACTGCTGTACCATGATATCTGTCCAATTTTTGGTTTCTTTTATTCTGTCAACGTCATTCTGCGCCAAGTGCATGACTATGCTACGTAATCTTGAAAGGTTTTTGTAACAT

180

GTCATCACCACCATTCTGAACCTTCAAGAGCCAATTCATCTTTGTCTACTTTTGCAATAAGACTATGGAGCAAAATATTTGTATTACAATCAACACCAAGTAAGTCATTTTATTAAC

300

TAAATTGACTGTGAAAGTCGAATAACCCCTCGGGATATCCTTTTTCTGGTCCAAAGGGCTTGAAAGGTAACCTTTTTATGGTCCAAAAACCACACGATTGTGCGCCCTTGGACCTGA

420

AGATAAAAGAGGAAATAAATACCGACAACAATGGAAAAACAATTCAATATTTGGAGAGTACAAGAGCAACTTTCTGAATTTCCAGTTCAAAAAACAATTGGAAATTAATTTAAAGGCT

540

CGAGACCAATATATAGAACAATTAATGATTAAAAAACAATGGATACACCTGACAACATCGCAGTTTATGATCTTAAGAAAAAAGGTCAAAAATATCATAGAAAATGTAACATT

660

TAAGGTTTAAAAATAGGTCTGTGTTAGACGTTTTAAGATTCTGTATATAGTAGTTGATAAACCTTACATCCGAAAAATGAAATGTGAAGCTGTTTTTATTAATAATAAAGGCC

780

ATG TGC GAC ATT GAT TTA AAA ACT TGT GAA AAA AGT TCC AAG TAT AAA AAA AAA GAT ATC GTT GAT CTT GGG AAA AAA TGT GGT GTT G

900

met CYS ASP ILE ASP LEU LYS THR CYS GLU LYS SER SER LYS TYR LYS LYS LYS ASP ILE VAL ASP LEU GLY LYS LYS CYS GLY VAL A

1020

CCA TAT TTA CCC AAT GGT CGC GAA AAA ACT CGA CAA GTA ATA TGT ACT GAA ATC GTC AAT CAG TAC TCT ACA AAT CCT CCA TCG TCA T

1140

PRO TYR LEU PRO ASN GLY ARG GLU LYS THR ARG GLN VAL ILE CYS THR GLU ILE VAL ASN GLN TYR SER THR ASN PRO PRO SER SER S

1260

TCT GAA AGT GAC GAT GAC GAA GAT ATG GGA AAG CAA GAT CAA GAT CCA ATT TGT GGC ATA TCT GAA AAA GAA TGC AAT AAA TTG GAT A

1380

SER GLU SER ASP ASP ASP GLU ASP met GLY LYS GLN ASP GLN ASP PRO ILE CYS GLY ILE SER GLU LYS GLU CYS ASN LYS LEU ASP A

1500

GTT AAC CTA TTA GCA TTG GGA GAA TAT TGT GGT GTT GAT ATT TAT ACC CCT GAA GGC AAT CTA AAG ACG TCT AAA ACC ATA TGT AAA

1620

VAL ASN LEU LEU ALA LEU GLY GLU TYR CYS GLY VAL ASP ILE TYR THR PRO GLU GLY ASN LEU LYS THR SER LYS THR ILE CYS LYS S

1740

GTT AGT TTG AAA CAA AGT TCT CCA AAA GCT CAA AAA AGC CCG AAT AAA AAC GAT TTG ATC AAT ATG GAC CCG GCT AAC TTA AAA AAA

1860

VAL SER LEU LYS GLN SER SER PRO LYS ALA GLN LYS SER PRO ASN LYS ASN ASP LEU ILE ASN met ASP PRO ALA ASN LEU LYS LYS U

1980

GCT AAA AAA TTA GAT ATT GAT TGG AAA AAG AAA GAT ATA AAA GAA CTT CAA* TAT TTA ATT GCA AAA AAA TTA AAT TCT ATA AAG TCG C

2100

ALA LYS LYS LEU ASP ILE ASP TRP LYS LYS LYS ASP ILE LYS GLU LEU GLN TYR LEU ILE ALA LYS LYS LEU ASN SER ILE LYS SER V

2220

AGT CCA TCT CGT TCA CCA TCT CCA AGA AGG TCT CGT TCA CCA TCC CCT AAT ATT CAA AAT ATG AAG AAA AAG GAA CTT AAA GCT TTA C

2340

SER PRO SER ARG SER PRO SER PRO ARG ARG SER ARG SER PRO SER PRO ASN ILE GLN ASN met LYS LYS LYS GLU LEU LYS ALA LEU A

2460

AAG GCT TTA GGC GCA TCC GAT AAA AAA CTT GAT AAA ATG GAC AAA AAT CAA CTT ATT GAA TAT ATT TTA TCA CGA AAA AAA TCC CCA

2580

LYS ALA LEU GLY ALA SER ASP LYS LYS LEU ASP LYS met ASP LYS ASN GLN LEU ILE GLU TYR ILE LEU SER ARG LYS LYS SER PRO

2700

CCT GTT AGA AAG TCA AGA TCC CCA AGT CTT CGT CAG AGA TCA CGT TCT CCT GGC CGG TCG AGG TCC AAT AGC CCT GTT AGA CCA AAA

2820

PRO VAL ARG LYS SER ARG SER PRO SER LEU ARG GLN ARG SER ARG SER PRO GLY ARG SER ARG SER ASN SER PRO VAL ARG PRO LYS

2940

AGG GCT GCT GAT CTG TTA GGC AAA AAA GTA GTT GAA CTA AAA GAG TTG GCT AAA ATG GAG GGA CTA AAA AAA TGG AAG GAT AAA ACA

3060

ARG ALA ALA ASP LEU LEU GLY LYS LYS VAL VAL GLU LEU LYS GLU LEU ALA LYS met GLU GLY LEU LYS LYS TRP LYS ASP LYS THR

3180

TCC AAA ATG TTA AAA CAA GAT TTG GTT GAT TTT TTA TTA CAT GTT GGA GGT GGT GGA CAA TTA CCA AAA ACT CCA AGT CCT GTT AAA

3300

SER LYS met LEU LYS GLN ASP LEU VAL ASP PHE LEU LEU HIS VAL GLY GLY GLY GLY GLN LEU PRO LYS THR PRO SER PRO VAL LYS

3420

TCA GCT TCT CCT AAA ACT AGA GCT GAA CTA GCT CTA TTA AAA GTA CCA GAA CTG AAA GCT ATG GCT TTA AGT TAT GGT TTG AAA AAT

3540

SER ALA SER PRO LYS THR ARG ALA GLU LEU ALA LEU LEU LYS VAL PRO GLU LEU LYS ALA met ALA LEU SER TYR GLY LEU LYS ASN

3660

AAA GAT AAA ACT CCA TCA CAA TTA CGG AAA AAT GAT TTT ATT GAT TTT ATC ATA TTA TCG ACT CAA AAA TCA CAG TCT CCT CCT CCT

3780

LYS ASP LYS THR PRO SER GLN LEU ARG LYS ASN ASP PHE ILE ASP PHE ILE ILE LEU SER THR GLN LYS SER GLN SER PRO PRO PRO

3900

ATT CGG TCT AAG CCA AGG TCA AGT TCT CCC AAA TAT AAA ACT AAA GAC TTA GTT ACA ACT GAT GAA AGT GAT GGC GAA ATT GTG GTT

4020

ILE ARG SER LYS PRO ARG SER SER SER PRO LYS TYR LYS THR LYS ASP LEU VAL THR THR ASP GLU SER ASP GLY GLU ILE VAL VAL

425 430 435 440 445 450
 AAG ATC ACG CGA AAA CCT AAA TCT CCT AGA CGT TCT CCT CCA GCA TCT GTT AGA AGA TCA AGA ACA CCC TCG GTA CCA AAG TCA CCT T
 LYS ILE THR ARG LYS PRO LYS SER PRO ARG ARG SER PRO PRO ALA SER VAL ARG ARG SER ARG THR PRO SER VAL PRO LYS SER PRO SE
 455 460 465 470 475 480
 GCG AGA CCA AGA TCC AAG TCA CCT TCT GTT AGA GCA GAA ATT ACC GAC GAT GAA GGT GAA ACA CCA CCT TCA TCA GTT AGA CCC AAG T
 ALA ARG PRO ARG SER LYS SER PRO SER VAL ARG ALA GLU ILE THR ASP ASP GLU GLY GLU THR PRO PRO SER SER VAL ARG PRO LYS SE
 485 490 495 500 505 510
 CCT TCT GTT AGA CCC AAG TCA CCA TCA GTT AGA CCA AGA TCC AAG TCA CCT TCA GTT AGA CCC AAG TCA CCT TCT GCG AGA CCA AGA T
 PRO SER VAL ARG PRO LYS SER PRO SER VAL ARG PRO ARG SER LYS SER PRO SER VAL ARG PRO LYS SER PRO SER ALA ARG PRO ARG SE
 515 520 525 530 535 540
 AAG TCA CCA TCA GTT AGA TCC AAG TCA CCA TCA GTT AGA CCC AAG TCA CCA TCA GTT AGA CCC AAG TCA CCA TCA GTT AGA CCA AGA T
 LYS SER PRO SER VAL ARG SER LYS SER PRO SER VAL ARG PRO LYS SER PRO SEP VAL ARG PRO LYS SER PRO SER VAL ARG PRO ARG SE
 545 550 555 560 565 570
 AAG TCA CCT TCT GTT AGA CCC AAG TCA CCT TCA GTT AGA CCA CCT CCA CCA GAT CCT TCT AAC TCT CCG AGT ATA ACA GTC GAC CCC T
 LYS SER PRO SER VAL ARG PRO LYS SER PRO SER VAL ARG PRO PRO PRO ASP PRO SER ASN SER PRO SER ILE THR VAL ASP PRO SE
 575 580 585 590 595 600
 GTA ACT CCT CCA TCA TCT GTT AAA ATA AAA AAG AGA CCA GAG TTA CCC GAG TAT AAA GGC GGA AAT GCG GAT AGA GAT TTG GAA ATC C
 VAL THR PRO PRO SER SER VAL LYS ILE LYS LYS ARG PRO GLU LEU PRO GLU TYR LYS GLY GLY ASN ALA ASP ARG ASP LEU GLU ILE L
 605 610 615 620 625 630
 GCT CGT AGA CGT GGC TAT AAA GTT ATA CCG GTA AAG GGA GAT GGA AAC TGT CTT TTC AGA GCT GTT GGG AAA TCT TTG AGA TTG AAC C
 ALA ARG ARG ARG GLY TYR LYS VAL ILE PRO VAL LYS GLY ASP GLY ASN CYS LEU PHE ARG ALA VAL GLY LYS SER LEU ARG LEU ASN G
 635 640 645 650 655 660
 AAT ATA AAA TAT TCT CAC GAA GAT TTA AGA GCA CAG GTT GTA ACT TAT CTT ACA TCT CAC AAA GAA TTT TTA GAA CCA TAT CTC GAA T
 ASN ILE LYS TYR SER HIS GLU ASP LEU ARG ALA GLN VAL VAL THR TYR LEU THR SER HIS LYS GLU PHE LEU GLU PRO TYR LEU GLU T
 665 670 675 680 685 690
 GTT ACT GAA AGT GGT GAT ACT ACA CCA CAA GAA TAT GCA AAA AAT GTT GAA AGA TAT ATA AAA AAT ATA TCT AAA CCT GGA ACA TGG G
 VAL THR GLU SER GLY ASP THR THR PRO GLN GLU TYR ALA LYS ASN VAL GLU ARG TYR ILE LYS ASN ILE SER LYS PRO GLY THR TRP G
 695 700 705 710 715 720
 GAT TTT ATT TGC TTA AGA GTG TTG AGT GAA ATA TTG AAG GTT AAA TTT AAC CTT TTA ATT TTA AAT ACA AGA AAT TTT CAA GTT ATA T
 ASP PHE ILE CYS LEU ARG VAL LEU SER GLU ILE LEU LYS VAL LYS PHE ASN LEU LEU ILE LEU ASN THR ARG ASN PHE GLN VAL ILE S
 725 730 735 740 745 750
 AAT AAT GAT ACT TTT AAA CCT TTA ATA CCT CTT GGT TTT ATA GAT GAC TAT CAT TAT ACA GCA TTG ACA CCT TTG TAT GCA GAA CCA A
 ASN ASN ASP THR PHE LYS PRO LEU ILE PRO LEU GLY PHE ILE ASP ASP TYR HIS TYR THR ALA LEU THR PRO LEU TYR ALA GLU PRO I
 755 760 765 770 775 780
 GCT GTT TTA GAA AAT GAA ACG CCT ACA CCT TCA ATT GCT CCA TCC ATT AGA CCC CCT CCA TCC ATC ATA CCT TCG ATT GCA CCT TCG G
 ALA VAL LEU GLU ASN GLU THR PRO THR PRO SER ILE ALA PRO SER ILE ARG PRO PRO PRO SER ILE ILE PRO SER ILE ALA PRO SER V
 785 790 795 800 805 810
 AGA CCA TCT ATT GCA CCT TCG GTT AGA CCA TCT ATT GTA CCA GCA ACT CCG TCT ATT GTA CCA TCC CTC AAA CCT TCT GTT GTA CCA A
 ARG PRO SER ILE ALA PRO SER VAL ARG PRO SER ILE VAL PRO ALA THR PRO SER ILE VAL PRO SER LEU LYS PRO SER VAL VAL PRO L
 815 820 825 830 835 840
 TTA ACA CCA TCT ATT GCA CCT AAA ATT CCC CCT CCA GTT TTT GGT CCT GTT AAA CCA TTA TCA TCA ACT AGA GAA TTG TTG GAA GAA G
 LEU THR PRO SER ILE ALA PRO LYS ILE PRO PRO PRO VAL PHE GLY PRO VAL LYS PRO LEU SER SER THR ARG GLU LEU LEU GLU GLU AL
 845 850 855 860 865 870
 GAT AAA GTT CAC CCA TAT GTA AGA GAA GAT ATA TCG CAG CTG GCA CGA GCA GAA GAA CAA ATT TTA GTT AGC CTT GGT ATG TAA
 ASP LYS VAL HIS PRO TYR VAL ARG GLU ASP ILE SER GLN LEU ALA ARG ALA GLU GLU GLN ILE LEU VAL SER LEU GLY met ***
 poly A 3380 3440
 TCATAAAAAATGATTTTAAACCTAAAAATATTAATTAATAAATGATGGAATCTAAGCATTAAACAATATTATTCATGATAAAATTAATTAATTTGATTGATTTTGTCTTGAGAAAAA
 3500 3560
 ATCTCTGATAGGATTAATAAATTAACCGTAAAAACAAAGAGTTGAATCTCGGTATTATTCACAAAGCCACTCAAAAAGTTCTTAAGACCATAACCGATAAAAAACGGTTATTGAGT
 3620 3680
 AAAAAGTCAATTTAACAATTATATCTTAGTTATGGAAGAAGACGATCCAAGTTTGAAGATATTAATAAATAACAAATTTGTAATTGATATAGATAATCAAAACCATAATTGGTACGGA
 3740 3800
 TTTAAGAGGAGATGTTGAACCTTTAACAAAAACATTGGTTGAAATTTGTGATAAATTTAAATTAATAACAACGTGCCTTTAAATTTGAATACAAGCGATGACTTGGATGAAGATAC
 3860 3920
 AAATTGCTAATGAGATACATGGATGGGACTAAATATGCAGAAAGTGAAAGTGATGATAATAATGAGGATTAATTTTTGTAGATAATAAATGCTTTTGAACGTAATAATTGGAGAT
 3980 4040
 AAGTCTATGGGAATTTGTGATGCTTTGGTAATCATTCACCTTTTGGTGGGTATCCTTCTTAACCTTATCAATAGTAAGTAATATATATTAATAAATTTTTTAATTTCTATAGAAGTT
 4100 4160
 AAATTAAACGATTGAAAGTTAATAAATGAAATACGAAGATTTTATAAACATTTGTCTTTACAAGTAAAAAATTTGGCCTTGATTGGTCCAAGTATATCAATCAAGTTAGGCTCTAC
 4220 4284
 ACCTAAAAATTCATACCAAGTTAGAGGTACCATTAATTTGAATTAAGATTGTTATTTTTCTATCAAAATTTATGGTGAAACTATCCG

B

2 of /de 2

MicroD

Figure 7. Location of other open reading frames in the region of the TIV genome near the L96 gene.

(A) The 4254 bp sequenced region was analyzed for open reading frames (ORFs) equal to or greater than 60 codons in length in all three possible reading frames on each strand. Note that ORFs F and N extend into the unsequenced flanking regions. (B) Predicted molecular weights of the putative proteins encoded by ORFs A to N translating from the beginning of the ORF to the termination codon. Again note that the true weight of the putative proteins encoded by ORFs F and N are greater than indicated in the table as these ORFs extend for an undetermined number of codons out of the sequenced region.

A

B

ORF ¹	NUMBER OF AMINO ACIDS	PREDICTED MOLECULAR WEIGHT	pI
A	68	7,944	11.85
B	74	8,929	10.31
C	62	7,069	9.28
D	157	18,583	6.22
E	149	17,246	7.00
F	61	6,984 (incomplete)	9.83
G	60	7,257	10.55
H	104	9,949	4.58
I	66	7,523	11.85
J	223	22,459	4.30
K	102	10,761	9.28
L	75	9,545	11.20
M	73	8,504	10.42
N	79	8,996 (incomplete)	5.92

¹ See Appendix C for the precise location of the ORFS in the sequenced region.

of direct repeat sequences places them just past the mid-point of the L96 ORF. In contrast, only two inverted repeat sequences greater than 12 nucleotides in length were found. Three inverted repeats were observed to be located within the 300 bp preceeding the putative L96 start codon (Figure 6A).

3. Protein sequence

The primary ORF described in the previous section was observed to encode a putative protein of 867 amino acids (Figure 6A) with a predicted size of 96 kDa and an estimated pI of 10.91. On the basis of this molecular information the gene was given the name L96. L to represent the fact that the gene is expressed late in TIV infection, and 96 to represent the predicted size of the putative gene product. Analysis of the predicted amino acid sequence revealed serine, proline, and lysine to be the most predominant amino acids comprising 13.4, 12.2, and 11.9%, respectively, of the putative protein (Table V). Using the method of Kyte and Doolittle (1982) a hydropathy plot was constructed for the putative L96 gene product (Figure 6B). This plot showed a largely hydrophilic protein with a small hydrophobic region near the C-terminus.

D. Construction of expression clones

In order to increase our knowledge of the L96 gene and its gene product it was decided to attempt to express the L96 gene in *Escherichia coli*. Ideally these experiments would provide a means of purifying the L96 gene product so that it may be studied in greater detail and its function elucidated. The gene product most likely cannot be easily purified from TIV-infected insect cells because all of the high molecular weight proteins are very minor components of TIV virions (Krell and Lee, 1974). Also, if the L96 gene product, or a derivative of it, can be produced in *E. coli* and then purified, it could be used to make an antibody specific for L96. This antibody could in turn be used to detect the native protein in TIV-infected insect cells, and aid in both the purification of this protein and the elucidation of its biological function.

Table V. Amino acid composition of the L96 gene product.

	amino acid	number of residues	percent composition
NON-POLAR:	Ala	37	4.268
	Ile	49	5.652
	Leu	71	8.189
	Met	9	1.038
	Phe	12	1.384
	Pro	106	12.226
	Trp	3	0.346
	Val	57	6.574
POLAR:	Asn	31	3.576
	Cys	10	1.153
	Gln	21	2.422
	Gly	31	3.576
	Ser	116	13.379
	Thr	36	4.152
	Tyr	22	2.537
BASIC:	Arg	60	6.920
	His	5	0.557
	Lys	103	11.880
ACIDIC:	Asp	46	5.306
	Glu	42	4.844

1. pUC 19 based constructs

a) Unfused construct

The most direct way of expressing the L96 gene in *E. coli* was to simply subclone the gene into a suitable vector. In this construct the L96 gene is under the control of its native TIV promoter, and therefore expression requires that the *E. coli* RNA polymerase recognize this promoter.

In order to facilitate construction of the expression clones, the 3.25 kbp BglII/HindIII fragment of L10 (Figure 8) containing the major portion of the L96 ORF was subcloned into plasmid pUC 19. The subcloning was performed by codigesting L10 with BglII, EcoRI, HindIII, and PstI and then ligating the ensuing mixture of DNA fragments to pUC 19 digested with BamHI and HindIII. BamHI and BglII yield fragments of DNA with identical 5' cohesive termini which, when ligated together, produces a hybrid site that is not recognized by either enzyme. Of the possible recombinant clones produced by this ligation only one was 5.9 kbp in size and yielded the correct banding pattern after restriction endonuclease analysis. This clone was named W7 (Figure 8). The TIV insert in this clone may be easily isolated by electroelution of the 3.25 kbp fragment produced by codigestion of W7 with EcoRI and HindIII. The pUC 19 multiple cloning site (MCS) has an EcoRI site just outside the BamHI site used in the subcloning and there are no internal EcoRI sites in this insert. Thus this clone provides a ready source of the major portion of the L96 ORF.

Subcloning of the L96 gene into pUC 19 could not be done by simple excision of a fragment of L10 containing the gene and ligation of this fragment to the vector as all enzymes suitable for the excision step cut within the L96 ORF. Clone W7 contains the 3' end of the L96 ORF; hence, all that is required is that the 5' end of the ORF be inserted in the HindIII site of this clone as illustrated in Figure 9. The 5' end of the L96 ORF was located on a 2.5 kbp HindIII fragment of L10. Thus, L10 DNA was digested with HindIII and the 2.5 kbp band isolated by electroelution. Ligation of this fragment to HindIII digested W7 DNA yielded the desired clone (Figure 9). This clone was named W8, and was distinguished from

Figure 8. Subcloning of the major portion of the L96 gene.

To provide a ready source of the major portion of the L96 ORF the 3.25 kbp BglII/HindIII fragment of L10 was subcloned using the BamHI/HindIII sites of plasmid pUC 19. L10 was digested with BglII, EcoRI, HindIII, and PstI thereby producing only one fragment with BglII/HindIII ends. Ligation of this mixture to BamHI/HindIII codigested pUC 19 yields only one possible recombinant clone (W7). The ligation of BglII to BamHI termini is explained in the text, and the brackets around the BglII site in W7 denote that this site was destroyed in the subcloning process..

B = BamHI

B⁰ = null BamHI site

Bg = BglII

E = EcoRI

H = HindIII

Hp = HpaI

K = KpnI

S = Sall

Figure 9. Subcloning of the L96 gene into plasmid pUC 19.

The L96 gene, under the control of its native TIV promoter, was subcloned into pUC 19. Clone W7 contained the entire L96 ORF except for 0.63 kbp from the 5' end of the ORF. The missing region of the ORF was located on the 2.5 kbp HindIII fragment of L10. Hence, the L96 ORF and native promoter was reconstructed by ligation of electroeluted 2.5 kbp HindIII fragment of L10 DNA to HindIII linearized W7 DNA. The resulting clone containing the L96 ORF under the control of its native promoter was named W8.

B = BamHI

B⁰ = null BamHI site

Bg = BglII

E = EcoRI

H = HindIII

Hp = HpaI

S = SalI

HindIII

HindIII

1kbp

L96 ORF
Start

LIGATE

1kbp

W8

the one with the 2.5 kbp HindIII fragment in the opposite orientation by restriction endonuclease analysis.

b) 5' fusion construct

In general eukaryotic genes are frequently poorly expressed and the gene products are rapidly degraded in *E. coli* cells. TIV is no exception in this regard as heterologous expression in *E. coli* cells of either the L96 gene (as visualized by Coomassie Blue staining of SDS-PAGE gels, see section III.E.1.a.) or the major TIV capsid protein gene (as visualized by Western analysis, Tajbakhsh, 1988) when encoded by plasmids possessing these genes under the control of their native promoters has not been detected. Hence, it was decided to attempt to express the L96 gene as a hybrid protein fused to a bacterial polypeptide. The particular *E. coli* polypeptide selected was β -galactosidase as this approach has been used successfully on several occasions (Davis *et al.*, 1981; Slocombe *et al.*, 1982; Casadaban *et al.*, 1983; Weinstock *et al.*, 1983; Lathe *et al.*, 1984; Offensperger *et al.*, 1985; Carrol and Laughon, 1987; Rutledge *et al.*, 1988). This construct is useful even when the foreign gene is fused only to the *lac* promoter as has been shown by the heterologous expression of the TIV capsid gene in *E. coli* (Tajbakhsh, 1988)

To circumvent the concern that the bacterial RNA polymerase would not recognize the TIV promoter preceeding the L96 gene and therefore prevent expression of the gene, it was decided to fuse the 5' end of the L96 ORF to the sequences encoding the N-terminus of the 107 amino acid alpha peptide of β -galactosidase located on plasmid pUC 19. The MCS of pUC 19 contains several unique restriction endonuclease sites at the 5' end of the alpha peptide sequence which can be used to create the fusion (Figure 8). In creating the fusion construct it is desirable to have as little as possible of the L96 5' untranslated sequences. Fortunately an HaeIII site (GGCC) is located 5 nucleotides upstream of the L96 translational start codon (Figure 6A). Unfortunately there are other HaeIII sites within the L96 ORF; hence, the construction of the fusion clone containing the entire L96 ORF is not easy to do.

The fusion of the L96 ORF to the alpha peptide of pUC 19 required the remaining 0.63 kbp of the ORF located on the 2.5 kbp HindIII fragment of L10 (Figure 10A). Hence, L10 DNA was digested with HindIII and the 2.5 kbp band was isolated by electroelution. This fragment was found to encode a single HaeIII site, 5 base pairs upstream of the L96 ORF. Therefore, the electroeluted 2.5 kbp HindIII fragment of L10 was digested with HaeIII to yield fragments 0.63 and 1.8 kbp in size, and the 0.63 kbp fragment was isolated by electroelution.

The fusion construct was prepared by ligating the 0.63 kbp HaeIII/HindIII fragment described above along with the 3.25 kbp HindIII/EcoRI insert from clone W7 to pUC 19. The vector was prepared for ligation by digestion with HindIII followed by filling in of the recessed 3' ends using the Klenow fragment of DNA Polymerase I, and then digestion with EcoRI. This yielded a linearized pUC 19 molecule with one blunt terminus and one EcoRI terminus. Since HaeIII leaves a blunt end there is only one possible recombinant product from this ligation, namely that consisting of the L96 ORF fused at its 5' end to the 5' end of the alpha peptide sequences of pUC 19 (Figure 10A). The orientation of the insert was verified by restriction endonuclease analysis, and the fusion verified by DNA sequencing (Figure 10B). This clone was named W10.

c) "Sandwich" constructs

The fusion of a foreign polypeptide to β -galactosidase appears to significantly enhance the stability of that polypeptide in the bacterial cell, and fusion proteins are sometimes more easily purified by chromatographic procedures established for β -galactosidase (Carroll and Laughon, 1987). However, the fusion protein in clone W10 has only 6 amino acids which are derived from the alpha peptide of β -galactosidase (Figure 10B); hence, any such benefits are likely to be nonexistent. The fusion protein produced by W10 contains only 6 β -galactosidase derived amino acids as it was constructed such that transcription was initiated from the alpha peptide promoter, but translation stops at the L96 translational stop codon. In order to make use of the advantages of fusion to β -galactosidase described previously, it was decided to fuse

Figure 10. Fusion of codons 1 to 867 of the L96 ORF to codons 1 to 6 of the alpha peptide of β -galactosidase of pUC19.

(A) Construction of clone W10 with the L96 ORF fused to the alpha peptide of pUC 19. The L96 ORF was reconstructed by isolation of the 3.25 kbp insert from W7 containing the major portion of the L96 ORF and the 0.63 kbp *Hae*III/*Hind*III fragment from the 2.5 kbp *Hind*III fragment of L10, and ligation of these two fragments to pUC 19 which had been *Hind*III digested, Klenow treated to fill in the recessed 3' termini, and *Eco*RI digested. This resulted in clone W10 possessing the 5' end of the L96 ORF fused to the 5' end of the alpha peptide.

(B) Verification of the fusion by DNA sequencing. The unbracketed numbers indicate *lacZ* amino acids, the bracketed numbers indicate pUC19 MCS amino acids, the unnumbered amino acids indicate the junction region, and the bold numbers indicate L96 amino acids.

B = BamHI

B⁰ = null BamHI site

Bg = BglII

E = EcoRI

H = HindIII

Ha = *Hae*III

(Ha) = *Hae*III site not regenerated in the subcloning

Hp = *Hpa*I

S = *Sal*I

A

B

<i>LacZ</i>										<i>L96</i>	
ATG	ACC	ATG	ATT	ACG	CCA	AGC	TCC	AAT	ATG	TGC	TAA
Met	Thr	Met	Ile	Thr	Pro	Ser	Sc ₁	Asn	Met	Cys	 - - -
	1	2	3	4	(1)	(2)			1	2	

both the 3' and 5' ends of the L96 ORF to portions of the alpha peptide. The ensuing "sandwich" construct may produce a product that is stable in the *E. coli* cell. "Sandwich" type constructs have been successfully used to express *Herpes simplex* virus type 1 glycoprotein C in *E. coli* (Amann *et al.*, 1984). To increase the probability of success two different sandwich constructs were engineered: the first contained the entire L96 ORF except for 14 codons at the 3' end, while the second contained an 859 bp *Sall*/*PvuII* fragment encoding a 287 amino acid peptide from the C-terminus of the putative L96 gene product (Figure 11).

The 0.86 kbp *Sall*/*PvuII* fragment was prepared by codigesting clone W7 (Figure 12A) with *Sall* and *PvuII*, and then isolating the 0.86 kbp band by electroelution. The *Sall* terminus of this fragment was blunt ended by filling in the recessed 3' end via Klenow treatment. The fragment was then cloned into pUC 19 which had been codigested with *EcoRI* and *SmaI*, and the *EcoRI* terminus blunt ended by Klenow treatment. The ligation was possible as both *PvuII* and *SmaI* produce blunt ends. Clearly the TIV insert can enter the vector in either orientation; hence, the desired clone was selected from amongst the various recombinant clones via restriction endonuclease analysis. The fusions at both the 3' and 5' ends of the insert were verified by DNA sequencing, and the clone was named W18 (Figure 12B).

The sandwich construct containing most of the L96 ORF was prepared in a manner similar to that used to construct clone W10. As described previously the major portion of the L96 ORF was located on the 3.25 kbp insert in clone W7. In the construction of clone W10 described previously this 3.25 kbp insert was isolated by codigestion with *HindIII* and *EcoRI* followed by electroelution of the 3.25 kbp band. Digestion of this fragment with *PvuII*, which cuts 14 codon upstream of the natural translational stop codon, yields two fragments 1.35 and 1.9 kbp in size (Figure 12C). The 1.9 kbp band contains most of the L96 ORF and was isolated by electroelution. Ligation of this fragment with the 0.63 kbp *HaeIII*/*HindIII* fragment from the 2.5 kbp *HindIII* fragment of L10 (Figures 10A and 12C) yields the desired sandwich construct (Figure 12C). Again the desired construct was selected from amongst the

Figure 11. Regions of the L96 ORF selected for the construction of "sandwich" expression clones.

Two different fragments of the L96 ORF were selected for the engineering of the "sandwich" constructs. The first was a 0.86 kbp Sall/PvuII fragment from the 3' end which was used to construct clone W18 (Figure 12A); while the second encoded the entire ORF, except for 14 codons at the 3' end, and was used to construct clone W19 (Figure 12C).

Ha = HaeIII

S = Sall

P = PvuII

Figure 12. Construction of "sandwich" expression clones.

(A) Construction of clone W18 containing the 0.86 kbp PvuII/SalI fragment of the L96 ORF fused at both the 3' and 5' ends to the alpha peptide of β -galactosidase of pUC 19. W7 was codigested with PvuII and SalI, and the 0.86 kbp fragment isolated by electroelution. The SalI end was blunt ended by Klenow treatment, and the fragment ligated to EcoRI/SmaI codigested pUC 19 in which the recessed 3' terminus of the EcoRI end had been filled in by Klenow treatment. (B) DNA sequence verification of both fusions in clone W18. The bracketed numbers indicate pUC 19 MCS amino acids, the unnumbered amino acids indicate the junction region, and the bold numbers indicate L96 amino acids. (C) Construction of clone W19 containing the entire L96 ORF, except for 14 codons at the 3' end, fused at both the 3' and 5' ends to the alpha peptide of β -galactosidase of pUC 19. The 5' end of the ORF was isolated on a 0.63 kbp HaeIII/HindIII fragment as described in the construction of W10. The 3' end of the ORF was located on a 1.9 kbp HindIII/PvuII fragment isolated from clone W7 by isolation of the insert as described in the construction of W10, followed by PvuII digestion of this fragment and electroelution of the 1.9 kbp band. Ligation of these two fragments to pUC 19 prepared as described in (A) yielded W19. (D) DNA sequence verification of both fusions in clone W19. Amino acids are numbered as described in (B).

A = Aval

Bg = BglII

E = EcoRI

H = HindIII

Ha = HaeIII

(Ha) = HaeIII site not regenerated in the subcloning

K = KpnI

P = PvuII

S = SalI

(S) = SalI site not regenerated in the subcloning

A

B

Lac Z. (mes)	TIV										Lac Z
...CTA GAG GAT CCC CTC GAC.....	TCG CAG AAT TCA CTG...										
... Leu Glu Asp Pro Leu Asp	Ser	Gln	Asn	Ser	Leu	...					
(10) (11) (12) (13) 568	852	853	5	6	7						

C

MCS

D

Lac Z (mcs)							TIV	Lac Z				
...CTA GAG GAT CCC CCC AAT ATG.....							TCG CAG AAT TCA CTG...					
... Leu	Glu	Asp	Pro	Pro	Asn	Met	Ser	Gln	Asn	Ser	Leu	...
(10)	(11)	(12)	(13)			1	852	853	5	6	7	

recombinant clones by restriction endonuclease analysis, and both the 3' and 5' fusions were verified by DNA sequencing (Figure 12D). This clone was named W19.

2. Lambda gt11 based constructs

All of the fusions to β -galactosidase constructed thus far may be described as N-terminal fusions. In order to increase the chances of success attempts were made to construct C-terminal fusions to the β -galactosidase gene encoded by the lysogenic bacteriophage lambda gt11. Since this phage encodes the entire *lacZ* gene, as opposed to pUC 19 which encodes only the alpha peptide of β -galactosidase, these fusions may produce a product that is more stable in the *E. coli* cell than any of the N-terminal fusions described previously. Also, the fusion protein should possess enzymatic activity and should be recognized by anti- β -galactosidase antibody.

Lambda gt11 possesses a unique *EcoRI* site located at the 3' end of the β -galactosidase gene (Figure 13) which was used for the insertion of the TIV sequences. As in the construction of the "sandwich" clones two different constructs were engineered: the first contained the W18 insert encoding a 287 amino acid peptide from the C-terminus of the putative L96 gene product, while the second contained the W19 insert encoding the entire L96 ORF except for 14 codons at its 3' end. Both inserts were isolated by first digesting W18 and W19 DNA with *XbaI*, followed by partial Klenow treatment, Mung Bean Nuclease treatment, and *EcoRI* digestion. The resulting blunt/*EcoRI* ended inserts were then isolated by electroelution, and the *EcoRI* termini blunt ended by Klenow treatment. These L96 sequences were then ligated to lambda gt11 DNA which had been ligated to protect the *cos* site, *EcoRI* digested, and the *EcoRI* termini blunt ended by Klenow treatment. Ligation of DNAs prepared in this way should result in the in frame fusion of the L96 sequences to the β -galactosidase sequences. The ligated DNA samples were packaged *in vitro* to produce infectious phage particles. These were then plated on a bacterial lawn, and any recombinant bacteriophage were selected by plaque hybridizations using radiolabelled 0.86 kbp W18 insert as probe. Positive plaques were picked, replated, and rescreened. One strongly positive

Figure 13. Fusion of the L96 ORF to the β -galactosidase gene of lambda gt11.

(A) The lysogenic bacteriophage lambda gt11 possesses a lone EcoRI site, located at the 3' end of the sequences encoding the complete β -galactosidase gene, into which foreign DNA may be cloned. (B) Regions of the L96 ORF used for fusion to the β -galactosidase gene of lambda gt11.

A

λGT11

B

plaque in the second screening in each experiment was selected and used to prepare phage DNA. Restriction endonuclease analysis of these DNA samples showed no differences when compared to digested native lambda gt11 DNA.

The lack of incorporation of the L96 sequences into the lambda gt11 chromosome could possibly be due to the inserts not being blunt ended. As a result a simpler method of preparation of the L96 sequences for fusion to the β -galactosidase sequences was sought. The method selected was to isolate the 1.9 kbp HindIII/PvuII fragment of the L96 ORF. Straightforward Klenow fill in of both termini of this fragment, followed by ligation to lambda gt11 DNA (prepared as in the previous attempts except that the DNA was dephosphorylated with CIP after the Klenow treatment to fill in the EcoRI overhang) should result in the desired fusion. Again the ligated samples were packaged *in vitro* and plated. However, when the plaque hybridizations were performed the L96 probe was found to hybridize to native lambda gt11 DNA. Consequently, any positive plaques on the trial plates do not necessarily contain the TIV insert. As a result of this observation the approach of fusing the L96 ORF to the β -galactosidase gene of lambda gt11 was not pursued further as no method existed to select the recombinant clones.

3. pGEM-2 based constructs

The final expression clone which was constructed was based on plasmid pGEM-2 (Promega BioTec, Figure 14). This construct placed the L96 ORF directly behind the Bacteriophage T7 promoter located on this plasmid, and, since Bacteriophage T7 RNA Polymerase is commercially available, permitted *in vitro* transcriptional and translational analysis of the L96 ORF in addition to heterologous expression using the two plasmid system of Tabor and Richardson (1985). The clone was constructed by ligating the 0.63 kbp BamHI/HindIII fragment from clone W19 (Figure 14) containing the start of the L96 ORF and the 3.25 kbp HindIII/EcoRI TIV insert of clone W7 containing the major portion of the L96 ORF and its translational stop codon (Figure 14) to BamHI/EcoRI codigested pGEM-2.

Figure 14. Fusing of the L96 ORF to the Bacteriophage T7 RNA Polymerase Promoter.

The L96 ORF was fused to the T7 promoter of plasmid pGEM-2 via a three part ligation involving the 0.63 kbp BamHI/HindIII fragment of clone W19 containing the 5' end of the ORF, the 3.25 kbp HindIII/EcoRI insert of clone W7 containing the major portion of the ORF and its translation stop codon, and BamHI/EcoRI codigested pGEM-2. The resulting clone was named W16.

B = BamHI

(Bg) = BglII site not regenerated in the subcloning

E = EcoRI

H = HindIII

(Ha) = HaeIII site not regenerated in the subcloning

K = KpnI

S = Sall

Multiple Cloning Site

HindIII
PstI
SalI
AccI
HincII
XbaI
BamHI
RvaI
SmaI
SacI
EcoRI

pGEM-2

BamHI
EcoRI

E

S

K

H

3' END OF L96 ORF

LIGATE

E

L96 ORF

S

K

H

1 kbp

W16

W7

EcoRI
HindIII

W19

BamHI
HindIII

H

B

5' END OF L96 ORF

As shown in Figure 14 only one recombinant clone was possible. The construct was verified by restriction endonuclease analysis, and was named W16 (Figure 14).

Efficient translation initiation of mRNA in *E. coli* requires the ribosomes to recognize the ribosome binding site of about 40 nucleotides in length. This includes the Shine-Dalgarno region and its upstream sequence, the spacer region between the Shine-Dalgarno region and the initiation codon, and the beginning of the gene (Curry and Tomich, 1988). The Shine-Dalgarno sequence (GGAGG) is complementary to the 3' end of the 16S ribosomal RNA and may stabilize the initiation complex formed between the mRNA and the ribosome (Curry and Tomich, 1988). In addition, the distance between the Shine-Dalgarno region and the initiation ATG is important and can range from 5 to 10 nucleotides (Curry and Tomich, 1988). The two plasmid heterologous expression system of Tabor and Richardson (1985) for the expression of foreign genes placed under control of the Bacteriophage T7 RNA Polymerase promoter could be attempted using clone W16 directly; however, in light of the above discussion the introduction of a Shine-Dalgarno region, properly spaced from the translational start ATG, between the T7 promoter and the L96 ORF should greatly facilitate the heterologous expression of the L96 gene in *E. coli*.

Oligomeric cassettes containing a Portable Translational Initiation Site (PTIS) are commercially available (Pharmacia). The PTISs contain a five base Shine-Dalgarno region with a 3' adjacent poly(A) tract followed by an initiator ATG codon eight bases from the Shine-Dalgarno region (Figure 15). This combination of features optimizes the initiation of translation. Each PTIS also possesses 5' overlap restriction sequences at the ends which facilitate the ligation of the PTIS into a vector. Hence, the PTIS with BamHI termini was ligated to BamHI digested W16. This resulted in clone W17 which possessed the PTIS inserted between the T7 promoter and the L96 ORF (Figure 15). Moreover, as verified by DNA sequencing, the perfectly spaced ATG of the PTIS is in frame with the L96 ORF (Figure 15).

Figure 15. Introduction of a Shine-Dalgarno site 5' of the L96 gene fused to the Bacteriophage T7 RNA Polymerase Promoter.

W16 DNA was digested with BamHI and ligated to the BamHI Portable Translational Initiation Site (PTIS) supplied by Pharmacia. This resulted in the introduction of a Shine-Dalgarno site (GGAGG) ideally spaced (eight nucleotides) from a translational start codon (ATG) which was in frame with the L96 gene, all under the control of the T7 promoter. The resulting clone was named W17.

B = BamHI

E = EcoRI

E. Heterologous expression of the L96 gene in *E. coli*

1. pUC 19 based constructs

a) Unfused construct

The heterologous expression of a cloned foreign gene in *Escherichia coli*, and analysis of the protein products involves three basic steps: 1) growth of the bacterial cells harbouring the recombinant plasmid encoding the foreign gene; 2) harvesting and lysis of the bacterial cells; and, 3) SDS-PAGE analysis of the proteins released by lysis of the cells. If the foreign gene has been placed under the control of an inducible promoter, the growth step will obviously include the induction of the promoter by suitable means.

Clone W8 is a pUC 19 based clone possessing the L96 ORF under the control of its native promoter; hence, it can be used to investigate the heterologous expression of this gene in *E. coli*. The proteins from untransformed and both pUC 19 and W8 transformed *E. coli* HB101 cells were subjected to SDS-PAGE analysis, followed by Coomassie Blue staining to visualize the bands. Although the L96 ORF is not fused behind the *lacZ* promoter in W8, it does lie in the same orientation as the *lacZ* gene (Figure 9) hence both uninduced and IPTG induced cultures were analysed (for an explanation of the induction of the *lac* promoter with the gratuitous inducer IPTG see the next section). The data shown in Figure 16 reveal that no extra bands are visible in the recombinant construct lanes as compared to the control lanes regardless of whether or not the culture was induced.

b) 5' fusion construct

The fusion of the L96 ORF to the 5' end of the sequences encoding the alpha peptide of β -galactosidase on plasmid pUC 19 (clone W10) placed the TIV gene under the control of the *lac* promoter. This promoter can be induced by the addition of isopropylthio- β -D-galactoside (IPTG) to the growth media. Induction with the gratuitous inducer IPTG results in repressor protein inactivation and the induction of transcription and translation of the *lacZ* gene. Clone W10 contains the L96 ORF fused at the 5' end to the sequences coding for the N-terminus of the alpha peptide. In this construct 9 additional amino acids are added prior to

Figure 16. Expression study of the L96 gene in *E. coli* cells using constructs with the L96 ORF either fused to its native TIV promoter or to the *lac* promoter.

Plasmid W10, containing the L96 ORF fused at its 5' end to the 5' end of the alpha peptide sequences of plasmid pUC 19, and plasmid W8, containing the L96 ORF under the control of its native TIV promoter, were transformed into *E. coli* HB101 cells. Cells were grown, induced, and lysates were analyzed by SDS-PAGE (4% stack/10% base) and Coomassie Blue staining.

M = protein markers

a = pUC 19 transformed *E. coli* HB101

b = IPTG induced pUC 19 transformed *E. coli* HB101

c = W10 transformed *E. coli* HB101

d = IPTG induced W10 transformed *E. coli* HB101

e = untransformed *E. coli* HB101

f = IPTG induced untransformed *E. coli* HB101

g = W8 transformed *E. coli* HB101

h = IPTG induced W8 transformed *E. coli* HB101

M a b c d e f g h M

the putative initial L96 methionine residue, six of which are derived from the *lacZ* gene (Figure 10B). This clone was used for expression studies in *E. coli* HB101 cells. The proteins from untransformed and both pUC 19 and W10 transformed HB101 cells were subjected to SDS-PAGE analysis, followed by Coomassie Blue staining to visualize the bands. The cells were grown in both the presence and the absence of IPTG. The data shown in Figure 16 reveal that no extra bands are visible in the recombinant construct lanes as compared to the control lanes regardless of whether or not the culture was induced.

c) "Sandwich" constructs

The β -galactosidase gene is convenient for making translational fusions because it is possible to remove its translational initiation region along with up to at least the first 27 amino acid codons without affecting β -galactosidase enzymic activity (Casadaban *et al.*, 1983). An additional result of the translational fusion to β -galactosidase is the formation of hybrid proteins that can be used in protein studies. The hybrid protein can be readily identified by its β -galactosidase enzymatic activity (Casadaban *et al.*, 1983). Offensperger *et al.* (1985) reported the production of a functional β -galactosidase complex in JM83 cells harbouring a recombinant plasmid with an in frame fusion at both ends of the gene encoding the entire sequence of the pre-S2 region of human hepatitis B virus envelope protein to the alpha peptide sequences of pUC 8. Thus, efficient complementation of the *lacZ* Δ M15 gene product of JM83 and production of functional β -galactosidase does occur when foreign sequences are fused in a "sandwich" type arrangement at the 5' end of the sequences encoding the alpha peptide.

In light of the failure of the 5' fusions to show expression of the L96 gene (section III.E.1.b), it was decided to test the "sandwich" constructs from section III.D.1.c. for the production of functional β -galactosidase. Thus clones W18 and W19 were transformed into *E. coli* JM101 cells. These cells possess the *lacZ* Δ M15 deletion in the β -galactosidase gene and therefore do not produce functional β -galactosidase. When these cells are transformed by plasmids encoding the alpha peptide of β -galactosidase, and are induced by IPTG, they

produce functional enzyme. This may be visualized by growing the transformed cells on media containing IPTG and Bluo-Gal. The enzyme cleaves the Bluo-Gal and results in the production of a chromogenic compound which gives the bacterial colony a characteristic blue colour instead of the white coloured colonies produced by cells which do not contain functional β -galactosidase. *E. coli* JM101 transformed with either sandwich clone, W18 or W19, were observed to give rise to blue colonies while non-transformed JM101 cells gave rise to white colonies on IPTG/Bluo-Gal indicator plates. These results prompted the conclusion that the fusion protein was being synthesized, and that a functional enzyme was produced in both instances.

The positive result demonstrating the synthesis of the fusion protein lead to attempts to visualize this protein in SDS-PAGE analyses of lysates from induced JM101 cells harbouring the sandwich constructs. The method of Amann *et al.* (1984) was used, with modifications. JM101 harbouring plasmids W18 and W19 were grown to mid-log phase at which point the *lac* promoter was induced by the addition of IPTG. The incubation was continued for a further 2 hours prior to the harvesting the culture and the lysing of the cells. SDS-PAGE analysis of the lysates, followed by visualization of the bands by Coomassie Blue staining, revealed no differences between the recombinant and control lanes (Figure 17).

Since evidence existed showing that a fusion protein was indeed being produced, yet no extra bands were visible in SDS-PAGE analysis of total cellular proteins, it was decided to fractionate the bacterial culture into extracellular, and intracellular soluble and insoluble proteins. The assumption made in these experiments is that the fusion protein is indeed being produced, but cannot be detected because it is "hidden" by an abundant *E. coli* protein. Fractionation of the culture, particularly into intracellular soluble and insoluble proteins, will hopefully eliminate this problem and result in visualization of the fusion protein in SDS-PAGE analysis of induced cultures. The data from this experiment revealed the presence of two major differences among the various trials (Figure 18). The intracellular soluble fraction showed a band at 31 kDa in all transformed lanes which was not present in the

Figure 17. Expression study of the L96 gene in *E. coli* cells using "sandwich" constructs.

Plasmid W18, containing a 0.86 kbp fragment from the 3' end of the L96 ORF fused at both the 3' and 5' ends of the fragment to the sequences encoding the alpha peptide of β -galactosidase of plasmid pUC 19; and plasmid W19, containing the entire L96 ORF--except for 14 codons at its 3' end--similarly fused to the alpha peptide sequences, were transformed into *E. coli* strain JM101 cells. Cells were grown and induced, and lysates were analyzed by SDS-PAGE (4% stack/10% base) and Coomassie Blue staining.

M = protein markers

a = pUC 19 transformed *E. coli* JM101

b = W18 transformed *E. coli* JM101

c = W19 transformed *E. coli* JM101

M a b c M

Figure 18. Fractionation of the proteins from the expression studies of the L96 "sandwich" constructs in *E. coli* into intracellular soluble and insoluble, and extracellular protein fractions.

E. coli JM101 transformed with plasmids W18 or W19 (these plasmids are described in the legend to Figure 15) were grown, induced, and harvested. The harvested cells were fractionated into intracellular soluble and insoluble (A), and extracellular (B) fractions as described in the methods, and the fractions analyzed by SDS-PAGE (4% stack/10% base) and Coomassie Blue staining.

M = protein markers

a = untransformed *E. coli* JM101

b = pUC 19 transformed *E. coli* JM101

c = W18 transformed *E. coli* JM101

d = W19 transformed *E. coli* JM101

A

SOLUBLE

INSOLUBLE

M a b c d a b c d M

B

M a b c d M

untransformed JM101 lane. Clearly this indicates that the observed band is due to a pUC 19 gene, and not the fusion protein as it is present in the nonrecombinant pUC 19 lane. Most likely this band is β -lactamase produced by the ampicillin resistance gene encoded on pUC 19 (β -lactamase has been reported to migrate in the range 28-32 kDa). The second notable difference occurred in the intracellular insoluble fraction. All trials except the W19 transformed JM101 lane were observed to yield a band at 28 kDa. Since this band was present in the nontransformed, as well as some of the transformed, JM101 lanes it is difficult to attach any significance to this observation.

Heterologous expression of the major TIV capsid gene in *E. coli* when under the control of the *lac* promoter, but not fused to the 5' end of the *lacZ* gene, has been observed (Tajbakhsh, 1988). Expression of the TIV capsid gene under these conditions was only detected in a Western analysis of the proteins from lysed *E. coli* cells using anti-capsid antisera as the primary antibody. This type of analysis was not possible in the case of the L96 gene product as antibody against this protein was not available. However, antisera against β -galactosidase are commercially available; hence, a similar type of analysis could be performed using this antibody to probe a Western blot of the proteins from induced cultures harbouring the "sandwich" expression clones from section III.D.1.c.

The intracellular soluble and insoluble protein samples of induced JM101 cells harbouring either clone W18 or clone W19 were electrophoresed through an SDS-PAGE gel as described in the legend to Figure 18. After electrophoresis the gel was not stained, and the proteins were electroblotted onto nitrocellulose. The filter was then probed with a 1:10,000 dilution of rabbit anti- β -galactosidase IgG (supplied by 5 Prime--3 Prime Inc.) as primary antibody, followed by detection of the positive bands using goat anti-rabbit IgG-alkaline phosphatase conjugate (provided in the Protoblot Immunoblotting System supplied by Promega Biotec) as secondary antibody. When the colour reaction was performed on this filter no bands were observed except when the reaction was allowed to proceed for a long time (> 15 min, when the manufacturer states that positive bands should be visible within 1 to 15 min). Moreover,

at these longer reaction times the bands which did appear were very faint, and were identical in all lanes--even the untransformed JM101 lane (data not shown). The strongest band in the blot was approximately 100-120 kDa in size, which is in the range expected for the fragment of β -galactosidase encoded by the *E. coli* cell. These results suggested that the epitope the antibody recognizes is located on the portion of β -galactosidase supplied by the bacterium, and not on the alpha peptide.

In order to investigate the possibility that the antibody was only faintly detecting the plasmid encoded alpha peptide-L96 fusion fragment while strongly detecting the bacterium encoded major fragment of β -galactosidase, it was decided to transform the "sandwich" clone DNAs into *E. coli* cells which have the entire β -galactosidase gene deleted. Hence, any positive bands in Western analysis of the proteins from these cells using the anti- β -galactosidase antibody will be due to the alpha peptide-L96 fusion proteins. Thus, W18 and W19 DNAs were separately transformed into *E. coli* HB101 (which lacks the β -galactosidase gene). Cultures of the resulting cells were grown to mid-log phase and induced with IPTG. After two further hours of growth the cells were harvested, washed, and lysed by boiling in Laemmli Buffer. The resulting protein samples were run through an SDS-PAGE gel, electroblotted onto nitrocellulose, and probed with the anti- β -galactosidase antibody. Again bands were observed only after prolonged reaction times in the colour reaction. Moreover, the bands were again identical in all trials, including the untransformed HB101 lane, and were also identical to those observed in the Western blot of the proteins from JM101 cells harbouring these plasmids (data not shown). At this point it was suspected that the bands being observed were due to the secondary antibody preparation recognizing *E. coli* epitopes even though it had been pre-competed with an *E. coli* lysate. Hence, the experiment was repeated except that the step involving binding of the anti- β -galactosidase antibody (the primary antibody) was omitted. When the colour reaction was performed on this blot the same bands as had been observed in the two previous blots were detected (data not shown). Thus, it was concluded that the anti- β -galactosidase antibody preparation did not detect any

protein in these analyses, and that the bands observed were due to the antibody preparation recognizing other epitopes.

2. pGEM-2 based constructs

a) Unlabelled proteins

Maximal production of the L96 gene product when the gene was placed under the control of the Bacteriophage T7 RNA Polymerase promoter was accomplished using the protocol of Tabor and Richardson (1985). *E. coli* harbouring both pGP1-2 (the plasmid encoding Bacteriophage T7 RNA Polymerase) and W17 (the recombinant clone with the L96 ORF under the control of the T7 promoter) were grown in rich media at 37°C until the culture was in early stationary phase. The temperature was then shifted to 42°C thereby inactivating the lambda repressor on the Bacteriophage T7 RNA Polymerase gene of pGP1-2. This permitted transcription of the T7 polymerase gene, and resulted in production of the enzyme. After 25 min at the elevated temperature rifampicin was added to the culture to inhibit the host cell *E. coli* RNA polymerase, and the culture incubated at 37°C for a further 2 h to permit transcription and translation to occur. Under these experimental conditions the only RNA polymerase which is functional in the *E. coli* cell is the T7 polymerase. Due to the specificity of this enzyme for its promoter, and the fact that that L96 ORF is the only gene under control of the Bacteriophage T7 RNA Polymerase promoter in the cell, a significant enrichment of the amount of the L96 gene product produced in the cell should result. The cells were harvested and lysed by boiling in Laemmli buffer for 5 min. Electrophoretic analysis of the resulting protein samples on SDS-PAGE gels followed by Coomassie Blue staining revealed the presence of a faint extra band at approximately 165 kDa in the W17 lane (Figure 19). Allowing for errors in the estimation of the sizes of large proteins from SDS-PAGE analysis, this band is in the region which would be predicted for a dimer of the L96 gene product. The gel in Figure 19 also reveals the presence of a band at 29 kDa in the pGEM-2 lane which is not present in any of the other lanes. With reference to the structure of pGEM-2 and the relative orientations of the T7 promoter and the ampicillin resistance

Figure 19. Expression study of the L96 gene in *E. coli* cells using Bacteriophage T7 RNA Polymerase promoter elements.

Plasmid W17, which contained the L96 ORF and an ideally spaced Shine-Dalgarno site and translational start codon fused behind the Bacteriophage T7 RNA Polymerase Promoter of plasmid pGEM-2, was transformed into pGP1-2 transformed *E. coli* K38 cells. Plasmid pGP1-2 possesses the Bacteriophage T7 RNA Polymerase gene under the control of the temperature sensitive lambda cI-857 repressor. The cells were grown to early stationary phase at 37°C, and the repressor inactivated by shifting the temperature to 42°C for 25 min. Rifampicin was then added to inhibit the host cell RNA polymerase, and the temperature shifted back to 37°C for 2 h. The cells were harvested, lysed, and analysed by SDS-PAGE (4% stack/10% base) and Coomassie Blue staining.

M = protein markers

a = pGP1-2 transformed *E. coli* K38

b = W17 + pGP1-2 transformed *E. coli* K38

c = pGEM-2 + pGP1-2 transformed *E. coli* K38

M a b c M

gene (Figures 14 and 15), this band is most likely β -lactamase which has been overproduced due to the experimental conditions utilized. Evidence in support of this conclusion comes from the *in vitro* transcription and translation experiments with pGEM-2 described in section III.F. In these experiments pGEM-2 was observed to yield a transcript 2.1 kb in size which, with reference to the structure of pGEM-2, indicates that the transcript runs through the complete ampicillin resistance gene. Thus, all evidence indicates that the assignment of the 29 kDa band to β -lactamase is a valid one.

b) *In vivo* labelled proteins

Foreign genes placed under the control of the Bacteriophage T7 RNA Polymerase promoter can be selectively labelled *in vivo* using the two plasmid system of Tabor and Richardson (1985). *E. coli* K38 cells harbouring both pGP1-2 and W17 were grown to mid-log phase, and the cells from a 0.2 ml aliquot of this culture harvested. After washing, the cell pellet was resuspended in M9 media supplemented with thiamine and 18 amino acids (cysteine and methionine were excluded). Growth was induced by incubation of this culture at 30°C for 1 h, and the lambda repressor on the T7 polymerase gene of pGP1-2 inactivated by increasing the temperature to 42°C. As in the production of unlabelled proteins described in the previous section this resulted in transcription of the T7 polymerase gene and production of the enzyme. Rifampicin was then added to inhibit the host cell *E. coli* RNA Polymerase, and the cells pulsed with ³⁵S-methionine. Since the only active RNA polymerase within the bacterial cell should be the T7 polymerase, and due to the specificity of this polymerase for its promoter, only those genes under the control of the T7 promoter should be transcribed. As a result the only proteins synthesized are those whose genes are under the control of the T7 promoter; hence, there should be preferential labelling of these proteins. After pulsing the cells were harvested, lysed, and the labelled proteins subjected to SDS-PAGE analysis. Visualization of the labelled protein bands by fluorography revealed the presence of strong bands at 42.5 and 61 kDa in the recombinant lane which were not present in the pGEM-2 control lane (Figure 20). One weaker band at 52 kDa was also

Figure 20. *In vivo* labelling and expression of the L96 gene in *E. coli* cells using Bacteriophage T7 RNA Polymerase promoter elements.

E. coli K38 cells transformed with plasmids pGP1-2 and W17 were grown to mid-log phase in rich medium. The cells from a 0.2 ml aliquot were harvested, washed with M9 medium, and resuspended in M9 medium supplemented with thiamine and 18 amino acids (cysteine and methionine were excluded). The resulting culture was incubated at 30°C for 1 h, and the temperature sensitive repressor inactivated by shifting the temperature to 42°C for 15 min. The host cell RNA polymerase was inactivated by adding Rifampicin and incubating at 42°C for a further 20 min before shifting the temperature to 30°C for 20 min. The culture was pulsed with 15 μ Ci of 35 S-labelled L-methionine for 5 min before harvesting the cells. The harvested cells were lysed and analyzed by SDS-PAGE (4% stack/10% base) followed by fluorography.

M = protein markers

a = *in vivo* labelled proteins from pGP1-2 transformed *E. coli* K38

b = *in vivo* labelled proteins from pGP1-2 + W17 transformed *E. coli* K38

c = *in vivo* labelled proteins from pGP1-2 + pGEM-2 transformed *E. coli* K38

a b c M

observed (Figure 20). The 165 kDa band observed in the unlabelled experiments was not found in these experiments. The pGEM-2 lane showed primarily two bands, 27.5 and 31.5 kDa in size, being labelled. Similar results were observed by Tabor and Richardson (1985), and these two bands were attributed to β -lactamase.

F. *In vitro* analysis of the L96 gene

1. *In vitro* transcription

The cloning of the L96 ORF into pGEM-2 permitted transcriptional studies to be performed on this gene. The high promoter specificity of Bacteriophage T7 RNA Polymerase results in the exclusive synthesis of the transcript which originates at the T7 promoter in *in vitro* transcription experiments using this enzyme. Moreover, sufficient RNA is produced in these reactions to permit visualization of the transcript by EtBr staining of agarose gels.

In vitro transcription reactions using Bacteriophage T7 RNA Polymerase as enzyme and the four NTPs as substrates were performed on circular pGEM-2 and W17, and W17 linearized with EcoRI at the end of the TIV insert distal to the T7 promoter (Figure 15). The resulting transcripts were analyzed by electrophoresis through an agarose-formaldehyde gel, and the bands visualized by staining with EtBr (Figure 21). The circular pGEM-2 and W17 templates yielded transcripts 2.1 and 4.8 kb in size, respectively, while the EcoRI linearized W17 template produced a 1.85 kb transcript. The bands were proven to be RNA, and not the template DNA, by treatment with RNase A as this resulted in the disappearance of the bands (data not shown).

2. *In vitro* translation

The RNA produced *in vitro* from clone W17 was used to synthesize the corresponding protein *in vitro*. Since the *in vitro* transcription reaction exclusively produces one RNA, namely the one originating from the T7 promoter, *in vitro* translation of this RNA should produce only one protein product--specifically that encoded by the L96 ORF which is fused to the T7 promoter. RNA produced by *in vitro* transcription of W17 with T7 RNA

Figure 21. Analysis of *in vitro* transcription products made from the L96 gene using Bacteriophage T7 RNA Polymerase.

Plasmid W17 DNA (see the Figure 17 legend for a description of this plasmid), either supercoiled or linearized with EcoRI, and supercoiled pGEM-2 DNA, were transcribed *in vitro* using Bacteriophage T7 RNA Polymerase as enzyme and the four ribonucleotide triphosphates as monomers. The *in vitro* synthesized RNAs were analyzed by electrophoresis through an agarose-formaldehyde gel and the bands visualized by EtBr staining.

M = RNA markers

a = *in vitro* transcription of T7 RNA Polymerase control template DNA (Promega Biotec)

b = *in vitro* transcription of supercoiled W17 DNA

c = *in vitro* transcription of EcoRI linearized W17 DNA

d = *in vitro* transcription of supercoiled pGEM-2 DNA

polymerase was purified by phenol/chloroform extraction followed by ethanol precipitation. The resulting concentrated RNA samples were translated into ^{35}S -methionine containing peptide *in vitro* using Rabbit Reticulocyte Lysate. The products of the translation reactions were analyzed by SDS-PAGE followed by fluorography (Figure 22). Both the circular and EcoRI linearized W17 derived RNAs produced a major product of 62 kDa and a minor product of 93 kDa. The pGEM-2 lane shows that two bands, 26.5 and 31.5 kDa in size, were synthesized. These bands are very similar to the 27.5 and 31.5 kDa proteins found to be labelled in the *in vivo* labelling experiment with pGEM-2; hence, allowing for errors in estimating protein sizes from polyacrylamide gels, these two bands were attributed to β -lactamase. The results of all of the expression related experiments performed using W17 are summarized in Table VI.

Figure 22. Analysis of *in vitro* translation product produced from the L96 gene.

The RNAs produced by *in vitro* transcription of plasmids pGEM-2 and W17 (see the Figure 19 legend for the *in vitro* transcription of these plasmids) were translated *in vitro* using Rabbit Reticulocyte Lysate. The translation reactions were analyzed by SDS-PAGE (4% stack/10% base) and fluorography.

M = protein markers

a = Rabbit Reticulocyte lysate without added exogenous RNA

b = *in vitro* translation of RNA produced by *in vitro* transcription
of supercoiled W17 DNA

c = *in vitro* translation of RNA produced by *in vitro* transcription
of EcoRI linearized W17 DNA

d = *in vitro* translation of RNA produced by *in vitro* transcription
of supercoiled pGEM-2 DNA

M a b c d

Table VI. Protein products observed in expression studies using the L96 ORF fused to the Bacteriophage T7 RNA Polymerase promoter.

construct	proteins (kDa) observed in <u>heterologous expression in <i>E. coli</i></u>		proteins (kDa) observed in <i>in vitro</i> translation of <u><i>in vitro</i> synthesized RNA</u>
	unlabelled	labelled	labelled
W17	165	-	-
	-	-	93 (minor)
	-	61.0 (main)	62 (main)
	-	52.0 (minor)	-
	-	42.5 (main)	-
pGEM-2	-	31.5 (main)	31.5 (main)
	29	-	-
	-	27.5 (minor)	26.5 (minor)

IV. DISCUSSION

A. Restriction site mapping of the TIV genome

One of the ongoing objectives in this laboratory is to assemble a comprehensive restriction endonuclease map of the entire TIV genome (Tajbakhsh *et al.*, 1986; Tajbakhsh, 1988). The 30 kb region analyzed in section III.B.2 is the largest single block of the TIV genome mapped to date. Previously the 11.8 kbp genomic EcoRI fragment #7 and a 3 kbp region of the 25 to 30 kbp genomic EcoRI fragment #1 encoding the TIV capsid gene had been characterized (Tajbakhsh, 1988). However, it is unfortunate that there is no overlap between the map deduced here and the two maps reported by Tajbakhsh (1988).

As summarized in Table IV and Figures 2 and 3, the L96 gene is located on BamHI, BglII, EcoRI, PstI, XbaI, and XhoI genomic TIV fragments 18.65, 8.55, 8.75, >23.1, 20.75, and 12.75 kbp in size, respectively. Comparison of these results with the previously reported catalogue of restriction endonuclease banding patterns of genomic TIV DNA (Tajbakhsh *et al.*, 1986; Tajbakhsh, 1988) reveals that the L96 gene can be assigned to the published genomic TIV DNA fragments indicated in Table VII. However, due to the difficulty in accurately estimating the sizes of large DNA fragments from agarose gels, in some cases the two genomic bands nearest in size to that observed in this thesis are tabulated. In the cases where a submolar and a molar genomic fragment are of the same approximate size as the observed genomic fragment possessing the L96 gene, most likely the molar fragment is the one which encodes the L96 gene. This conclusion is based on the observation that the genomic fragments encoding the L96 gene yield bands of equal intensities when probed with the cloned gene.

Both of the L10 based probes were observed to select single bands in digests of genomic DNA with enzymes which do not cut within the probe sequence. The conclusion which may be drawn from these observations is that the L96 gene is present in only a single copy in the TIV genome. If the gene was present in more than one copy these results would not be possible.

Table VII. Possible TIV genomic DNA fragments encoding the L96 gene¹.

Restriction Endonuclease	Observed genomic fragment size(kbp)	Published ² genomic fragments of this size	Published ² size (kbp)
BamHI	18.65	#1, #2	19.0, 18.0
EcoRI	8.75	#9	8.7
PstI	>23.1	#1, #2	32.0, 26.0
XbaI	20.75	#1, #2	21.5, (20.0)
XhoI	12.75	#6, #7	(13.0), 12.5

¹ Values in parentheses represent submolar fragments.

² Tajbakhsh *et al.*, 1986; and Tajbakhsh, 1988

The problems encountered in the mapping of some of the adjacent restriction endonuclease sites in the deduction of the extended restriction endonuclease map in section III.B.2 are difficult to explain. In certain digests and codigests the predicted banding pattern is observed, and therefore the adjacent site may be mapped; while in others more fragments than predicted hybridized to the probes, and as a result these sites could not be mapped (eg. the 3' EcoRI site). One possible explanation for these observations is the presence of repeated sequences in the TIV genome (Tajbakhsh *et al.*, 1986). Specifically, if the sequence flanking the TIV insert in L10, particularly at the 3' end, is repeated at least twice in the genome then banding patterns such as those observed here would be predicted. The flanking sequence must be the region repeated as the insert in L10 does not appear to be repeated (see the previous paragraph). A second possible explanation is non-specific hybridization of the probe to the genomic DNA fragments due to the A-T richness of the TIV DNA in general (Thomas, 1961; Kawase and Hukuhara, 1967; Bellet and Inman, 1967; Tajbakhsh *et al.*, 1986), and of all TIV genes cloned to date in particular (Tajbakhsh, 1988; this thesis). This seems unlikely since both the hybridizations and the washes were performed under high stringency conditions (1% SDS, 1M NaCl, 65°C), although some evidence in support of nonspecific hybridization of L96 derived probes was observed in the screening of lambda gt11 based plaques (see section IV.B.2). A third possible explanation is incomplete digestion of the DNA by the enzyme(s). Again this is unlikely as the results are repeatable, and the banding patterns of the EtBr stained gels prior to blotting are identical to those previously reported for genomic TIV DNA (Tajbakhsh *et al.*, 1986; Tajbakhsh, 1988).

The null BamHI site found in L10 appears to be due to a deletion which occurred in the cloning process (Figure 3E). The evidence reported here indicates that the L96 gene, and therefore the TIV insert in L10, is located on either BamHI fragment #1 or #2. Thus, since the TIV insert in L10 is only 11.8 kbp in length while these two genomic fragments are 19.0 and 18.0 kbp in size, respectively, approximately 6 to 7 kbp have been deleted in the cloning process. Cloning and restriction endonuclease mapping of both of these genomic fragments,

followed by comparison of these maps to that of the TIV insert in L10 shown in Figure 2B, would confirm that a deletion occurred in the construction of L10.

B. Molecular characterization of the late TIV gene L96

1. Sequence characterization of the L96 gene

DNA sequencing of subcloned fragments of the TIV insert in L10 in the region to which the oligomeric probe #11-4 hybridized revealed the presence of a open reading frame 2604 nucleotides long. This ORF encoded a putative protein of 867 amino acids with a predicted size of 96 kDa and an estimated pI of 10.91. As mentioned in the previous section, the gene was found to be present in the TIV genome as a single copy, and had a transcript size of 3.6 kb as determined by Northern analysis (Tajbakhsh, 1988). The coding region of L96 was not found to be interrupted by introns. This observation was partly expected as introns have not been shown to be present in DNA viruses that replicated in the cytoplasm (Moss, 1985), and the TIV capsid gene sequence was shown to be intron-free (Tajbakhsh, 1988).

As discussed in the Introduction, cytoplasmic DNA viruses such as vaccinia virus and FV3 possess genes which lack typical prokaryotic or eukaryotic consensus regulatory sequences. However, analysis of the 5' untranslated region of the L96 gene revealed the presence of a CCAAT sequence and a putative TATA box 235 nt and 76 nt upstream from the putative ATG start codon, respectively. These motifs are similar to those reported for the TIV capsid gene (Tajbakhsh, 1988) where a CCAAT sequence and an A-T rich sequence were observed at 129 and 81 nt upstream, respectively, from the ATG start codon of the capsid gene (Table VIII). Of the two FV3 genes which have been sequenced to date (Willis *et al.*, 1984a; Beckman *et al.*, 1988), only one (ICR489) contained a TATA (AATAAAA at -41, 72 nt upstream from the ATG) and a CCAAT sequence (at -104, 135 upstream from the ATG) similar to eukaryotic motifs (Beckman *et al.*, 1988). The FV3 motifs are similar in appearance and location when compared with the A-T rich and CCAAT sequences of the TIV capsid gene and the TATA box and CCAAT sequence of the TIV L96 gene, although

TABLE VIII. Comparison of the FV3 ICR489 gene and the TIV L96 and capsid genes' 5' and 3' UTR regions.

Virus	Gene	5' UTR ¹		3' UTR	Reference
		CCAAT	TATA	poly A	
TIV	L96	-235	-76	+	This thesis Tajbakhsh, 1988
	capsid	-129	-81	+	
FV3	ICR489	-135	-72	-	Beckman <i>et al.</i> , 1988

¹ All distances are calculated in nucleotides upstream from the ATG start codon.

the L96 CCAAT sequence is slightly farther upstream (Table VIII). At least one eukaryotic polyadenylation signal (AATAAA) was located at the 3' end of the L96 ORF. A similar finding has been reported for the TIV capsid gene (Tajbakhsh, 1988); however, to date the question of the polyadenylation of TIV transcripts has not been resolved (Tajbakhsh, 1988; Tajbakhsh and Seligy, 1989).

Comparison of the codon usage patterns of the TIV capsid gene and the L96 gene revealed some significant differences. For example, the capsid gene utilizes both phenylalanine codons (TTT and TTC) with almost equal frequency, while the L96 gene utilizes primarily the TTT phenylalanine codon. Similar differences were also observed in the alanine, asparagine, lysine, threonine, and tyrosine codon usage patterns. These results are consistent with observations made by Ikemura (1985) who reported that the codon-choice patterns of individual genes from any one organism are often very different.

The putative protein predicted by the DNA sequence of the L96 ORF is relatively large in size (96 kDa). In a previous study on the size estimates of the polypeptides of TIV and TIV-infected hemocytes of *Galleria mellonella* larvae, Krell and Lee (1974) reported 28 viral proteins ranging from 17.5 to 300 kDa in size. At least one of these TIV specific proteins has an estimated molecular weight of 96 kDa. The putative L96 gene product is predicted to be highly basic and largely hydrophilic. One possible role for the L96 gene product may be viral DNA binding which generates the chromatin-like structure characteristic of *Iridoviridae* inner cores (Thomas and Williams, 1961; Klump *et al.*, 1983). The hydrophilic nature of the putative L96 protein is consistent with the properties of the coat protein which is involved in the formation of the highly hydrated *Iridoviridae* nucleoprotein cores (Klump *et al.*, 1983; Kelly, 1985). Protein sequence comparisons between L96 and protein sequences present in two sequence banks provided further evidence in support of this hypothesis. Specifically, this search showed that the L96 amino acid sequence had homology to the core protein (p15) and inner coat protein (p12) regions of the GAG polyproteins of both the AKV murine leukemia virus (38.8% identity in a 49 amino acid overlap) and the Friend spleen focus-forming virus

(33.0% identity in a 72 amino acid overlap). The regions of L96 which show these homologies to the GAG polyproteins overlap, and are located at the C-terminus of L96 (i.e. aa's 771 to 822 for the homology to the AKV murine leukemia virus GAG polyprotein, and aa's 771 to 848 for the homology to the Friend spleen focus-forming virus GAG polyprotein). These results support the hypothesis that L96 has a structural role in TIV since the regions of the GAG polyproteins to which it shows homology have roles in the formation of the internal structures of these viruses (Clark and Mak, 1983; Herr, 1984). Other lesser homologies were observed between L96 and the probable E2 protein of papillomavirus type 5, goose histone H5, the GAG polyprotein of feline sarcoma virus, the HB_eAg core antigen of hepatitis B virus, and the GAG-GPR80 prepeptide of feline leukemia virus. These observations also lend credence to the notion that L96 is involved in the formation of the inner TIV nucleoprotein core structure as both the papillomavirus E2 protein (Giri and Yaniv, 1988) and histone proteins (Lewin, 1983) are known to be DNA binding proteins, while the remaining proteins in the list are all GAG polyproteins or GAG polyprotein derived proteins and these are known to be involved in the formation of the inner structural features of various viruses (Lowy, 1985).

2. Expression of L96 in *E. coli* cells

The observation that a fragment of the TIV insert in L10 exclusively hybridizes to a 3.6 kb transcript in Northern analysis of total RNA from TIV-infected *Estigmene acrea* cells, followed by the DNA sequence evidence that this fragment of L10 is part of a large, 2.6 kbp open reading frame provided strong evidence that the L96 gene was in fact a true TIV gene which is expressed in virus infected cells. However, in order to obtain direct evidence that this is indeed the case, it is necessary to detect L96 gene product in TIV-infected insect cells. Unfortunately at present there is no method for the detection or means of purifying the L96 derived protein from these cells. The most obvious solution to this problem, and the approach followed here, was to attempt to express the cloned L96 gene in *E. coli* and to purify its gene product from lysed bacterial cells for use in production of L96 specific

antisera. Hopefully this antisera could be used to detect the L96 gene product in lysates of virus infected insect cells, and aid in the purification of this protein.

Initial expression studies in *E. coli* using the entire region of the TIV insert in L10 carrying the L96 gene failed to produce any measurable protein product, as assayed by searching for an "extra band" in Coomassie Blue stained SDS-PAGE gels of lysed *E. coli* cells harbouring the appropriate recombinant L96 expression plasmid, indicating that the L96 promoter was not recognized in *E. coli* cells. This result is not surprising as the failure of expression of viral genes in *E. coli* is actually quite common (Knipe, 1985), and similar findings were reported by Tajbakhsh (1988) using the TIV capsid gene and its promoter in *E. coli*. In these experiments the production, or lack thereof, of the TIV capsid protein in the *E. coli* cell was assayed by Western blot analysis of *E. coli* lysates using anti-capsid antisera as primary antibody. The visualization of the bands in the SDS-PAGE gels was done by Coomassie Blue staining, and not silver staining, because of the predicted nature of the L96 gene product. Specifically, a strong correlation between the intensity of response to Coomassie dyes and the basicity of a protein has been reported, with the number of dye molecules bound to the protein being approximately proportional to the number of positively charged amino acids it contains (Tal *et al.*, 1985). In this thesis it was found that the putative L96 gene product contained almost 20% basic amino acid residues (168 residues), with the predominant one being lysine (103 residues). Although the silver staining of proteins in gels is reported to be up to 100 times more sensitive than Coomassie Blue staining, the basis of the silver stain is not as well defined (Nielson and Brown, 1984; Chuba and Palchaudhuri, 1986). Additionally, Chuba and Palchaudhuri (1986) have shown that a strong correlation exists between increased staining with silver and the cysteine content of the protein. Since the putative L96 gene product has only 10 cysteine residues, as opposed to the 168 basic amino acid residues mentioned above, Coomassie Blue staining was the method of choice in this thesis.

Since TIV promoters do not appear to be recognized by *E. coli* RNA polymerase (Tajbakhsh, 1988; this thesis), it was decided to express the L96 gene in *E. coli* by fusing it to

a bacterial promoter and/or gene (Knipe, 1985). As discussed in sections III.D.1.b and c this approach has been used successfully on several occasions, most significantly in the heterologous expression of the TIV capsid gene in *E. coli* when fused to the *lacZ* promoter of plasmid pUC 19 (Tajbakhsh, 1988). A construct (W10) was subsequently created by fusing the 5' end of the L96 ORF to the sequences encoding the 9 N-terminal amino acids of the alpha peptide from plasmid pUC 19. In this construct transcription originates from the *lacZ* promoter, and translation begins at the initial methionine codon of the alpha peptide and should continue to the L96 termination codon. This construct was expressed in *E. coli* cells by induction of the *lacZ* promoter with the gratuitous inducer IPTG. SDS-PAGE fractionation of the expressed proteins followed by Coomassie Blue staining revealed the presence of no extra bands as compared to appropriate controls. Similar results were observed in the expression of the TIV capsid gene, when fused to the *lacZ* promoter, in *E. coli* (Tajbakhsh, 1988). However, in the latter experiments expression was detected by use of anti-capsid antibody as a probe after Western blotting of SDS-PAGE separated expressed proteins. Since no differences between the control and experimental lanes were observed in Coomassie Blue stained gels in these experiments, the level of capsid gene expression must have been relatively low.

One possible explanation for both the low level of expression of the capsid gene and the lack of any detectable expression of the L96 gene is rapid degradation of either or both of the mRNA and the polypeptide it encodes. Various workers have shown that these problems may be overcome by fusion of the foreign protein to a bacterial protein (Knipe, 1985), either in a 1:1 arrangement (Carroll and Laughon, 1987) or in a "sandwich" type arrangement in which the foreign protein is sandwiched between either the N- and C-termini of a single bacterial protein or the N-terminus of one bacterial protein and the C-terminus of another (Amann *et al.*, 1984). Thus, subsequent experiments involved the construction of recombinant plasmids containing portions of the L96 ORF fused at both the 5' and 3' ends to the sequences encoding the N- and C-termini, respectively, of the alpha peptide of

β -galactosidase encoded on plasmid pUC 19. The fusion of the L96 ORF to the sequences encoding the alpha peptide of β -galactosidase in these sandwich constructs was desirable for several reasons: 1) foreign proteins fused to β -galactosidase are frequently more easily purified from *E. coli* lysates using chromatographic procedures established for β -galactosidase than are unfused proteins (Carrol and Laughon, 1987); 2) the foreign gene is under the control of an inducible promoter; and, 3) expression of the fusion protein may be verified by enzymic assay *in vivo* (Casabadian *et al.*, 1983; Offensperger *et al.*, 1985).

In order to increase the chances of success two sandwich constructs were created. The first contained approximately one-third of the L96 ORF from the 3' end of the ORF (W18), while the second contained nearly the entire ORF (W19). Prior to SDS-PAGE analysis of lysates of *E. coli* cells harbouring these expression constructs, production of the fusion protein *in vivo* could be assayed by looking for the production of functional β -galactosidase in the *E. coli* cell. The alpha peptide encoded by pUC 19 complements the defective β -galactosidase produced by *E. coli* strains possessing the *lacZ* Δ M15 genotype, and results in the production of a functional enzyme. Production of functional β -galactosidase by *E. coli* may be visualized by growing the cells on media containing IPTG and Blueo-Gal. The gratuitous inducer IPTG results in transcription of the alpha peptide, while Blueo-Gal is a colourless compound which is cleaved to a blue one by the action of β -galactosidase. Both W18 and W19, when transformed into *E. coli* strain JM101 cells and plated on IPTG/Blueo-Gal indicator plates, were observed to yield blue colonies while untransformed JM101 cells yielded white colonies. These results indicated that a fusion protein was likely being produced, since the alpha complementation had occurred and a functional β -galactosidase was produced. However, it must be recalled that although transcription starts at the 5' end of the fused sequences, translation does not necessarily begin in this region (i.e. internal translational initiation) as the N-terminus of the alpha peptide is not required for complementation (Zabin and Fowler, 1978).

Since the evidence suggested that a fusion protein was being produced in each case, the proteins from induced *E. coli* cells harbouring one or the other of these constructs were subjected to SDS-PAGE separation followed by Coomassie Blue visualization. However, even with fractionation of the induced cultures into intracellular soluble and insoluble, and extracellular proteins, no additional bands were observed. These two results, production of functional β -galactosidase in the enzyme assay and no extra band in SDS-PAGE analysis of *E. coli* lysates, are difficult to reconcile because one indicates production of the fusion protein while the other doesn't. One possible explanation is that proteolysis of the fusion protein occurred. It is now well documented that mutationally altered proteins may be rapidly degraded in bacteria (Zabin and Fowler, 1978); thus, since only trace amounts of functional β -galactosidase are required for the production of a blue colony, what may be occurring is production of the fusion protein followed by rapid proteolytic degradation within the *E. coli* cell. Evidence in support of this theory comes from the expression of the TIV capsid gene in *E. coli* in which several polypeptides smaller than the full capsid protein were observed to react positively with anti-capsid antisera in Western analysis of the expressed proteins from lysed *E. coli* cells (Tajbakhsh, 1988). The production of blue colonies observed in this thesis may be explained by protection of a small amount of the fusion protein from proteolysis by its reaction with the Δ M15 fragment of β -galactosidase supplied by the bacterium in the formation of the functional enzyme. Thus, functional β -galactosidase is produced, but insufficient fusion protein is present to be visualized by Coomassie blue staining of SDS-PAGE gels. A second possible explanation for these results is that the fusion protein band is a minor one which is hidden behind an abundant *E. coli* protein band of equivalent size. If this is indeed the case the use of a two dimensional gel should reveal the fusion protein band. A third possible explanation is internal translational initiation within the fusion protein. Recently the production of N-terminally truncated proteins attributable to translational initiation at downstream AUG codons in the expression of several eukaryotic genes in *E. coli* have been reported (Preikesch *et al.*, 1988). These authors reported that downstream

translational initiation occurred at AUG codons located between 4 and 20 nucleotides 3' of sequences resembling bacterial Shine-Dalgarno sites (GGAGG). However, when the L96 ORF was searched for Shine-Dalgarno-like sequences, only two were found. Of these, only one possessed an ATG codon between 4 and 20 nucleotides downstream, but this codon was not in the L96-*lacZ* reading frame, and the ensuing ORF continued for only 3 codons before reaching a stop codon.

The results thus far support the hypothesis that the fusion protein is produced in the *E. coli* cells, but that only a small amount of it is present at any one time. This small amount is sufficient to produce alpha complementation, but insufficient to be detected by SDS-PAGE analysis. One method of testing for the production of the fusion protein was Western blot analysis of the proteins from lysed *E. coli* cells. As no antibody existed against the L96 region of the fusion protein, one against the β -galactosidase region was used. Unfortunately, no differences were observed between control and trial lanes in this experiment. The most likely explanation for these results is that the epitopes recognized by this polyclonal antibody are located on the portion of β -galactosidase supplied by the bacterium, and not on the alpha peptide. The use of an anti-alpha peptide antibody, if available, would eliminate this problem.

In order to circumvent these problems attempts were made to fuse the L96 ORF to the sequences encoding the complete β -galactosidase gene of the lysogenic bacteriophage lambda gt11. Three different constructs were attempted; however, in each case no recombinant bacteriophage were obtained. Moreover, it was observed that radiolabelled probes from both the middle and the 3' end of the L96 ORF hybridized non-specifically to lambda gt11 DNA. As a result the experiment was discontinued as no method existed for selecting the recombinant phage from the wild type ones.

The final approach taken to express the L96 gene in *E. coli* was to utilize the two plasmid system of Tabor and Richardson (1985). In this system the gene of interest was fused to the Bacteriophage T7 RNA Polymerase promoter, and the ensuing construct transformed into an

E. coli strain already harbouring a plasmid encoding the T7 RNA polymerase. The specificity of T7 polymerase for its own promoter, combined with the ability to inhibit selectively the host RNA polymerase with rifampicin, permits the exclusive expression of genes under the control of the T7 RNA polymerase promoter. Recently, this system has been used to express the Herpes Simplex Virus Type 1 DNA Polymerase gene in *E. coli* (Dorsky and Crumpacker, 1988).

A construct (W17) was engineered with the L96 ORF fused to the T7 promoter of plasmid pGEM-2. In order to enhance translational initiation within the *E. coli* cell an oligomeric cassette containing a Shine-Dalgarno or consensus ribosome binding site (GGAGG, Curry and Tomich, 1988) and a properly spaced 3' ATG codon was inserted between the T7 promoter and the L96 ORF such that the ATG codon was in frame with the L96 ORF. This construct, when transformed into *E. coli* cells already harbouring the T7 RNA polymerase encoding plasmid pGP1-2, was used in the production of both unlabelled and labelled proteins.

Isolation of unlabelled proteins from these cells after inhibition of the host RNA polymerase with rifampicin revealed only one difference between the experimental and control samples. Specifically, a very faint band was observed at approximately 165 kDa in the W17 sample. It is difficult to attach much significance to this observation as the band in question is considerably larger than that predicted by the DNA sequence of the L96 ORF, and it is extremely faint. The most likely explanation is that this band is either a gel artifact or a dimer of a lower molecular weight protein which was not completely dissociated by the treatment with SDS prior to electrophoresis.

More definitive results were obtained in the production of labelled proteins by pulsing the rifampicin inhibited cultures with ^{35}S -labelled methionine. Due to the rapid degradation of *E. coli* mRNAs (Tabor and Richardson, 1985) and the promoter specificity of T7 RNA polymerase (Tabor and Richardson, 1985), preferential transcription and translation of the L96 ORF and selective labelling of the protein products of the L96 ORF should occur. The

labelled proteins produced (42.5, 52.0 [minor], and 61.0 kDa in size) are smaller than the size predicted by the DNA sequence data for the putative L96 protein (96 kDa). The presence of several inverted repeats in the regions of the L96 ORF where transcripts encoding proteins of this size would terminate were observed. The presence of these inverted repeats likely rendered this region unfavorable for continuous transcription from the T7 promoter by acting as premature termination sites (Platt and Bear, 1983). These observations may also be used to hypothesize as to the reason for the lack of detectable fusion protein production in the β -galactosidase alpha peptide fusion experiments. Specifically, if only a small percentage of all initiated transcripts run through this region of the L96 ORF, then only sufficient complete fusion protein to produce complementation, and therefore blue colonies, is present and insufficient amounts of the fusion protein are present to be visualized in SDS-PAGE analysis.

In light of the difficulties encountered in all of these systems in obtaining sufficient protein for use as the immunogen in the production of an anti-L96 antibody, thought must be given to an alternative method of obtaining the desired antibody. One obvious approach would be the use of a synthetic peptide of the predicted L96 protein as immunogen in the production of an anti-L96 antibody. Unfortunately this approach goes beyond the scope and time frame of this study.

3. *In vitro* analysis of the L96 gene

One advantage of cloning the L96 gene into the T7 promoter containing vector pGEM-2 is that it permitted *in vitro* transcriptional and translational analysis of the L96 ORF. This approach has recently been used to obtain enzymatically active Herpes Simplex Virus Type 1 DNA polymerase (Dorsky and Crumpacker, 1988). *In vitro* transcription reactions using circular W17 DNA as template produced a 4.8 kb transcript. Since the TIV insert in this clone is approximately 4 kbp, this indicates that transcription was initiated at the T7 promoter, ran completely through the TIV insert in W17, and returned to plasmid sequences before terminating. Thus, if W17 DNA linearized with EcoRI at the end of the TIV insert in W17 distal to the T7 promoter was to be used as template in the *in vitro* transcription

reaction, "run off" transcription should occur and a transcript of approximately 4 kb should be produced. Experimentally a 1.85 kb transcript was produced, indicating that transcription terminates within the L96 ORF. One possible explanation for the smaller than expected size of this transcript is that linearization of the template DNA permits this DNA to form secondary structure which renders the L96 ORF unfavorable for transcription. These results provide some evidence in support of the premature transcriptional termination explanation for the sizes of the labelled proteins produced in the *in vivo* labelling experiment. The 4.8 kb transcript migrates to approximately the same location as the template W17 DNA in the agarose-formaldehyde gel thereby possibly indicating that the band is DNA and not RNA. This band was shown to be RNA, and not template DNA, by the fact that it disappears upon treatment of the sample with RNase A while similarly treated W17 DNA is unaffected.

In vitro translation of both the 4.8 and 1.85 kb *in vitro* synthesized RNAs produced identical proteins. The fact that the circular and linear W17 templates yield different *in vitro* transcripts, but identical proteins when these transcripts are translated *in vitro* is difficult to explain other than to suggest that the observed protein products are due to the translation of minor RNAs which are not visible in the EtBr stained agarose-formaldehyde gel of the transcripts. Both the *in vivo* labelling and the coupled *in vitro* transcription and translation produce a major band at approximately 61 to 62 kDa. This result could be due to incomplete transcription of the L96 ORF as described previously, or incomplete translation of the transcript--possibly due a requirement for another viral protein for complete translation of the mRNA.

C. Conclusions

In conclusion, my thesis work has led to the elucidation of the restriction endonuclease map of approximately 15% of the large TIV genome. This map should prove useful in obtaining the restriction endonuclease map of the complete TIV genome. This thesis has also provided considerable molecular information on TIV DNA sequence in general and on a second TIV gene in particular, and has permitted sequence comparisons to the TIV capsid

gene sequence and revealed the presence of several possible consensus regulatory sequences. The L96 gene was shown to be transcribed and translated both *in vitro* and *in vivo*, and to be heterologously expressed in *E. coli*. The availability of this cloned gene in several expression constructs should allow for the purification of the L96 gene product and subsequent antibody production and structural and functional analyses to be performed. Finally, the studies carried out here have increased our knowledge of the molecular organization of TIV, and have made it amenable to further in depth molecular analysis.

V. REFERENCES

- Arif, B. M., and W. Doerfler. (1984). "Identification and localization of reiterated sequences in the *Choristoneura fumiferana* MNPV genome." *EMBO J.* 3:525-529.
- Amann, E., M. Broker, and F. Wurm. (1984). "Expression of Herpes simplex virus type 1 glycoprotein C antigens in *Escherichia coli*." *Gene.* 32:203-215.
- Aubertin, A. M., L. Tondre, J. P. Martin, and A. Kirn. (1980). "Structural polypeptides of frog virus 3, phosphorylated proteins." *FEBS Letters.* 112:223-238.
- Baird, S. (1989). "The molecular isolation and characterization of endo- β -1,4-glucanase genes from *Bacillus polymyxa* and *Bacillus circulans*." M.Sc. Thesis, Biology Department, University of Ottawa, Ottawa, Canada.
- Balbas, P., X. Sorberon, E. Merino, M. Zurita, H. Lomeli, F. Valle, N. Flores, and F. Bolivar. (1986). "Plasmid vector pBR322 and its special-purpose derivatives--a review." *Gene.* 50:3-40.
- Beckman, W., T. N. Tham, A. M. Aubertin, and D. B. Willis. (1988). "Structure and regulation of the immediate-early frog virus 3 gene that encodes ICR489." *J. Virol.* 62:1271-1277.
- Bellett, A. J. D., and R. B. Inman. (1967). "Some properties of deoxyribonucleic acid preparations from *Chilo*, *Sericesthis* and *Tipula* iridescent viruses." *J. Mol. Biol.* 25:425-432.
- Ben-Ze'ev, A., and Y. Aloni. (1983). "Processing of SV40 RNA is associated with the nuclear matrix and is not followed by accumulation of low molecular weight RNA products." *Virology.* 125:475-479.
- Ben-Ze'ev, A., R. Abdulafi, and S. Bratosin. (1983). "Herpes simplex virus and protein transport are associated with the cytoskeletal framework and the nuclear matrix in infected BSC-1 cells." *Virology.* 129:501-507.
- Bertin, J., M. Frosch, and P. E. Lee. (1987). "Formation and maintenance of viroplasmic centers in *Tipula* iridescent virus-infected mosquito cells with deranged cytoskeletons." *Euro. J. Cell. Biol.* 43:215-222.
- Bibor-Hardy, V., M. Pouchelet, E. St. Pierre, M. Herzberg, and R. Simard. (1982). "The nuclear matrix is involved in herpes simplex virogenesis." *Virology.* 121:296-306.
- Bickle, T. (1982). "The ATP-dependent restriction endonucleases." in *The Nucleases*, edited by S. Linn and R. J. Robersts. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory, pp. 85-108.
- Biggin, M. D., T. J. Gibson, and G. F. Hong. (1983). "Puffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination." *Proc. Natl. Acad. Sci. USA.* 80:3963-3965.
- Bird, F. T. (1961). "The development of *Tipula* iridescent virus in the crane fly, *Tipula paludosa* Meig., and the wax moth, *Galleria mellonella* L." *Can. J. Microbiol.* 7:827-830.

- Bladon, T., M. Frosch, P. M. Sabour, and P. E. Lee. (1986). "Association of nuclear matrix proteins with cytoplasmic assembly sites of *Tipula* iridescent virus." *Virology*. 155:524-533.
- Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, J. L. Heynecker, H. W. Boyer, J. H. Crosa, and S. Falkow. (1977). "Construction and characterization of new cloning vehicles. II. A multipurpose cloning system." *Gene*. 2:95-113.
- Carlson, A., and J. Messing. (1984). "Efficiency in cloning and sequencing using the single-stranded bacteriophage M13." *J. Biotech.* 1:253-264.
- Carrol, S. B., and A. L. Laughon. (1987). "Production and purification of polyclonal antibodies to the foreign segment of β -galactosidase fusion proteins." in *DNA Cloning: a practical approach*. Volume 3, edited by D. M. Glover. Washington D. C.: IRL Press, pp89-111.
- Carter, J. B. (1973). "The mode of transmission of *Tipula* iridescent virus. I. Source of infection." *J. Invert. Path.* 21:123-130.
- Carter, J. B. (1974). "*Tipula* iridescent virus infection in the developmental stages of *Tipula oleracea*." *J. Invert. Path.* 24:271-281.
- Casadaban, M. J., A. Martinez-Arias, S. K. Shapira, and J. Chou. (1983). " β -galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast." *Methods in Enzymology*. 100:293-308.
- Caspar, D. L. D., and A. Klug. (1962). "Physical principles in the construction of regular viruses." *Cold Spring Harbor Symp. Quant. Biol.* 27:1-24.
- Chinchar, V. G., and A. Grilloff. (1984). "Isolation and characterization of a frog virus 3 variant resistant to phosphonoacetate: genetic evidence for a virus-specific DNA polymerase." *Virology*. 138:357-361.
- Chuba, P. J., and S. Palchaudhuri. (1986). "Requirement for cysteine in the colour silver staining of proteins in polyacrylamide gels." *Anal. Biochem.* 156:136-139.
- Clark, S. P., and T. W. Mak. (1983). "Complete nucleotide sequence of an infectious clone of Friend spleen focus-forming provirus: gp55 is an envelope fusion glycoprotein." *Proc. Natl. Acad. Sci. USA*. 80:5037-5041.
- Cuillet, M., F. Tripier, J. Braunwald, and B. Jacrot. (1979). "A low resolution structure of frog virus 3." *Virology*. 99:277-285.
- Curry, K. A., and C. S. C. Tomich. (1988). "Effect of ribosome binding site on gene expression in *Escherichia coli*." *DNA*. 7:173-179.
- Dale, R. M. K., B. A. McCiure, and J. P. Houchins. (1985). "A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA." *Plasmid*. 13:31-40.
- Darai, G., K. Anders, H. G. Koch, H. Delius, H. Gelderblom, C. Samalecos, and R. M. Flugel. (1983). "Analysis of the genome of fish lymphocystis disease virus isolated directly from epidermal tumors of pleurcnectes." *Virology*. 126:466-479.

- Darcy-Tripier, F., M. V. Nermut, J. Braunwald, and L. D. Williams. (1984). "The organization of frog virus 3 as revealed by freeze-etching." *Virology*. 138:287-299.
- Davis, A. R., D. P. Nayak, M. Ueda, A. L. Hiti, D. Dowbenko, and D. G. Kleid. (1981). "Expression of antigenic determinants of the hemagglutinin gene of a human influenza virus in *Escherichia coli*." *Proc. Natl. Acad. Sci. USA*. 78:5376-5380.
- DeBlois, R. W., E. E. Uzgis, D. H. Cluxton, and H. M. Mazzone. (1978). "Comparative measurements of size and polydispersity of several insect viruses." *Analyt. Biochem.* 90:271-288.
- Deilus, H., G. Darai, and R. M. Flugel. (1984). "DNA analysis of insect iridescent virus 6: Evidence for circular permutation and terminal redundancy." *J. Virol.* 49:609-614.
- Devauchelle, G., J. Attlias, C. Monnier, S. Barray, M. Cerutti, J. Guerilon, and N. Orange-Balange. (1985). "*Chilo* iridescent virus." *Curr. Top. Microbiol. Immunol.* 116:37-48.
- Dorsky, D. I., and C. S. Crumpacker. (1988). "Expression of Herpes Simplex Virus Type 1 DNA Polymerase gene by *in vitro* translation and effects of gene deletions on activity." *J. Virol.* 62:3224-3232.
- Dretzen, G., M. Bellard, P. Sassone-Corsi, and P. Chambon. (1981). "A reliable method for the recovery of DNA fragments from agarose and acrylamide gels." *Anal. Biochem.* 112:295-298.
- Dulbecco, R., and H. S. Ginberg. (1980). *Virology*. New York: Harper & Row, Publishers Inc.
- Escribano, J. M., E. Tabares. (1987). "Proteins specified by African swine fever virus: V. Identification of immediate-early, early, and late proteins." *Arch. Virol.* 92:221-232.
- Essani, K., R. Goorha, and A. Granoff. (1987). "Mutation in a DNA-binding protein reveals an association between DNA-methyltransferase activity and a 26,000-Da polypeptide in frog virus 3-infected cells." *Virology*. 161:211-217.
- Flugel, R. M. (1985). "Lymphocystis disease virus." *Curr. Top. Microbiol. Immunol.* 116:133-150.
- Franke, C. A., and D. E. Hruby. (1986). "Vaccinia virus: The hitchhiker syndrome." *VIth Poxvirus/Iridovirus Workshop*. Sept. 24-28, Cold Spring Harbor Laboratory.
- Giri, I., and M. Yaniv. (1988). "Study of the E2 gene product of the cotton rabbit papillomavirus reveals a common mechanism of transactivation among papillomaviruses." *J. Virol.* 62:1573-1581.
- Glitz, D. G., G. J. Hills, and C. F. Rivers. (1968). "A comparison of the *Tipula* and *Sericesthis* iridescent viruses." *J. Gen. Virol.* 3:209-220.
- Goorha, R. (1981). "Frog virus 3 requires RNA polymerase II for its replication." *J. Virol.* 37:496-499.
- Goorha, R. (1982). "Frog virus 3 DNA replication occurs in two stages." *J. Virol.* 43:519-528.

- Goorha, R. G., A. Granoff, and R. Tirey. (1978). "Macromolecular synthesis in cells infected by frog virus 3. VIII. The nucleus is the site of frog virus 3 DNA and RNA synthesis." *Virology*. **84**:32-50.
- Goorha, R. G., D. B. Willis, and A. Granoff. (1977). "Macromolecular synthesis in cells infected by frog virus 3. VI. Frog virus 3 replication is dependent on the cell nucleus." *J. Virol.* **21**:802-805.
- Goorha, R., and A. Granoff. (1979). "Icosahedral cytoplasmic deoxyriboviruses." in *Comprehensive Virology*, edited by R. R. Wagner and H. Fraenkel-Conrat. New York: Plenum Press. **14**:347-399.
- Goorha, R., and K. G. Murti. (1982). "The genome of frog virus 3, an animal DNA virus, is circularly permuted and terminally redundant." *Proc. Natl. Acad. Sci. USA*. **79**:248-252.
- Hall, D. W. (1985). "Pathobiology of invertebrate icosahedral cytoplasmic deoxyriboviruses (*Irioviridae*)." in *Viral Insecticides for Biological Control*, edited by K. Maramorosch and K. E. Sherman. Montreal: Academic Press Inc., pp.163-196.
- Hannahan, D. (1985). "Techniques for Transformation of *E. coli*." in *DNA Cloning--Volume 1*, edited by D. M. Glover. Washington, D.C.: IRL Press, pp. 109-135.
- Hefford, M. A., R. Brousseau, G. Prefontaine, Z. Hanna, J. A. Condie, and P. C. K. Lau. (1987). "Sequence of a lepidopteran toxin gene of *Bacillus thuringiensis* subsp *Kurstaki* NRD-12." *J. Biotech.* **6**:307-322.
- Hellman, D. M., J. R. Feramisco, J. C. Fiddes, G. P. Thomas, and S. H. Hughes. (1983). "Identification of clones that encode chicken tropomyosin by direct immunological screening of a cDNA expression library." *Proc. Natl. Acad. Sci. USA*. **80**:31-35.
- Herr, W. (1984). "Nucleotide sequence of AKV murine leukemia virus." *J. Virol.* **49**:471-478.
- Hills, G. J., and K. M. Smith. (1959). "Further studies on the isolation and crystallization of insect cytoplasmic viruses." *J. Insect Pathol.* **1**:121-128.
- Hohn, B. (1979). "In vitro packaging of lambda and cosmid DNA." *Methods in Enzymology*. **68**:299-309.
- Horwitz, M. S. (1985). "Adenoviruses and their replication." in *Virology*, edited by B. N. Fields. New York: Raven Press, pp. 433-476.
- Ikemura, T. (1985). "Codon usage and tRNA content in unicellular and multicellular organisms." *Mol. Biol. Evol.* **2**:13-34.
- Kalmakoff, J., and J. H. Tremaine. (1968). "Physiochemical properties of *Tipula* iridescent virus." *J. Virol.* **2**:738-744.
- Kawase, S., and T. Hukuhara. (1967). "Amino acid Composition of the *Tipula* iridescent virus." *J. Invert. Path.* **9**:273-274.
- Kelly, D. C. (1985). "Insect iridescent viruses." *Curr. Top. Microbiol. Immunol.* **116**:23-35.

- Kelly, D. C., and Tinsley. (1974). "Iridescent virus replication: a microscope study of *Aedes aegypti* and *Antheraea eucalypti* cells in culture infected with iridescent virus types 2 and 6." *Microbiologica*. 9:75-93.
- Klug, A., R. E. Franklin, and S. P. F. Humphreys-Owen. (1959). "The crystal structure of *Tipula* iridescent virus as determined by Bragg reflection of visible light." *Biochim. Biophys. Acta*. 32:203-219.
- Klump, A., J. Beaumais, and G. Devauchelle. (1983). "Structural and thermodynamic investigation of the *Chilo* iridescent virus (iridovirus type 6). *Arch. Virol.* 75:269-276.
- Knipe, D. (1985). "Molecular genetics of animal viruses." in *Virology*, edited by B. N. Fields. New York: Raven Press, pp. 129-143.
- Krell, P., and P. E. Lee. (1974). "Polypeptides in *Tipula* iridescent virus (TIV) and in TIV-infected hemocytes of *Galleria mellonella* (L.) larvae." *Virology*. 60:315-326.
- Kyte, J., and R. F. Doolittle. (1982). "A simple method for displaying the hydropathic character of a protein." *J. Mol. Biol.* 157:105-132.
- Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of Bacteriophage T4." *Nature*. 227:680-685.
- Lathe, R. F., M. P. Kieny, D. Schmitt, P. Curtis, and J. P. Lecocq. (1984). "M13 bacteriophage vectors for the expression of foreign proteins in *Escherichia coli*. The rabies glycoprotein." *J. Mol. App. Gen.* 2:331-342.
- Lee, P. E., and S. P. Brownrigg. (1982). "Effect of virus inactivation on *Tipula* iridescent virus-cell relationships." *J. Ultrastruct. Res.* 79:189-197.
- Lee-Chen, G. J., and E. G. Niles. (1988). "Transcription and translation mapping of the 13 genes in the vaccinia virus *HindIII* D fragment." *Virology*. 163:52-63.
- Lewin, B. (1983). *Genes II*. Toronto: John Wiley and Sons.
- Lowy, D. R. (1985). "Transformation and oncogenesis: retroviruses." in *Virology*, edited by B. N. Fields. New York: Raven Press, pp. 235-263.
- Manyakov, V. F. (1977). "Fine structure of the iridescent virus type I capsid." *J. Gen. Virol.* 36:73-79.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. (1982). *Molecular Cloning: a laboratory manual*. Cold Spring Harbor Laboratory.
- Maramorosch, K., and K. E. Sherman, editors. (1985). *Viral Insecticides for Biological Control*. Montreal: Academic Press.
- Martin, J. P., A. M. Aubertin, and A. Kirn. (1982). "Expression of frog virus 3 early genes after ultraviolet irradiation." *Virology*. 122:402-410.
- Mathieson, W. B., and P. E. Lee. (1981). "Cytology and autoradiography of *Tipula* iridescent virus infection of insect suspension cell cultures." *J. Ultrastruct. Res.* 74:59-68.

- Matthews, R. E. F. (1982). "Classification and nomenclature of viruses." *Intervirology*. 17:56-58.
- Meints, R. H., K. Lee, and J. L. Van Etten. (1986). "Assembly site of the virus PBCV-1 in a *Chlorella*-like green algae: Ultrastructural studies." *Virology*. 154:240-245.
- Mesnard, J. M., T. N. Tham, L. Tondre, A. M. Aubertin, and A. Kirn. (1988). "Organization of RNA transcripts from a 7.8 Kb region of the frog virus 3 genome." *Virology*. 165:122-133.
- Messing, J. (1983). "New M13 vectors for cloning." *Methods in Enzymology*. 101:20-78.
- Milavek, B., T. Hopkins-Davis, and C. M. Payne. (1984). "Identification of nuclear matrix proteins tightly bound to soluble Simian Virus 40 chromosomes." *Virology*. 134:405-420.
- Miller, H. (1987). "Practical aspects of preparing phage and plasmid DNA: growth, maintenance, and storage of bacteria." *Methods in Enzymology*. 152:145-170.
- Miller, R. C. (1975). "Replication and molecular recombination of T-phage." *Ann. Rev. Microbiol.* 29:355-376.
- Miner, J. N., S. L. Weinrich, and D. E. Hruby. (1988). "Molecular dissection of *cis*-acting regulatory elements from 5'-proximal regions of a vaccinia virus late gene cluster." *J. Virol.* 62:297-304.
- Moss, B. (1978). "Poxviruses." in *The molecular biology of animal viruses.*, edited by D. P. Nayak. New York: Marcel Dekker Inc. 2:849-890.
- Moss, B. (1985). "Replication of poxviruses." in *Virology*, edited by B. N. Fields. New York: Raven Press, pp. 685-703.
- Murti, K. G., and R. Goorha. (1983). "Interaction of frog virus 3 with the cytoskeleton. I. Altered organization of microtubules, intermediate filaments, and microfilaments." *J. Cell Biol.* 96:1248-1257.
- Murti, K. G., K. R. Porter, R. Goorha, M. Ulrich, and G. Wray. (1984). "Interaction of frog virus 3 with the cytomatrix. II. Structure and composition of the virus assembly site." *Exp. Cell Res.* 154:270-282.
- Murti, K. G., R. Goorha, and M. W. Klymkowsky. (1988). "A functional role for intermediate filaments in the formation of frog virus 3 assembly sites." *Virology*. 162:264-269.
- Nielson, B. L., and L. R. Brown. (1984). "The basis for coloured silver-protein complex formation in stained polyacrylamide gels." *Anal. Biochem.* 156:136-139.
- Offensperger, W., S. Wahl, A. R. Neurath, P. Price, N. Strick, S. B. H. Kent, J. K. Christman, and G. Acs. (1985). "Expression in *Escherichia coli* of a cloned DNA sequence encoding the pre-S2 region of hepatitis B virus." *Proc. Natl. Acad. Sci. USA.* 82:7540-7544.
- Ortin, J., and E. Vinuela. (1977). "Requirement of cell nucleus for African swine fever virus replication in VERO cells." *J. Virol.* 21:902-905.

- Platt, T., and D. Bear. (1983). "Role of RNA polymerase, rho factor, and ribosomes in transcription termination." in *Gene Function in Prokaryotes*, edited by J. Beckwith, J. Davies, and J. Gallant. New York: Cold Spring Harbor Laboratory, pp. 123-161.
- Preikisch, G., H. Ishihara, D. Tripiet, and M. Leinweber. (1988). "Unexpected translation initiation within the coding region of eukaryotic genes expressed in *Escherichia coli*." *Gene*. 72:179-186.
- Reed, K. C., and D. A. Mann. (1985). "Rapid transfer of DNA from agarose gels to nylon membranes." *Nucl. Acid Res.* 13:7207-7221.
- Roizman, B., and W. Batterson. (1985). "Herpesviruses and their replication." in *Virology*, edited by B. N. Fields. New York: Raven Press, pp. 497-526.
- Rosel, J. L., P. L. Earl, J. P. Weir, and B. Moss. (1986). "Conserved TAAATG sequence at the transcriptional and translational initiation sites of vaccinia virus late genes deduced by structural and functional analysis of the *HindIII* H genome fragment." *J. Virol.* 60:436-449.
- Russel, M., and P. M^c del. (1985). "Replacement of the *fip* gene of *Escherichia coli* by an inactive gene cloned on a plasmid." *J. Bact.* 159:1034-1039.
- Rutledge, R. G., J. M. Neelin, and V. L. Seligy. (1988). "Isolation and expression of cDNA clones coding for two sequence variants of *Xenopus laevis* histone H5." *Gene*. 70:117-126.
- Sanger, F., S. Nicklen, and A. R. Coulson. (1977). "DNA sequencing with chain-terminating inhibitors." *Proc. Natl. Acad. Sci. USA*. 74:5463-5467.
- Seagull, R. W., P. E. Lee, and M. Frosch. (1985). "Comparison of microtubules and microfilaments in *Tipula* iridescent virus-infected and uninfected cells." *Can. J. Biochem. Cell Biol.* 63:543-552.
- Skalka, A. M. (1977). "DNA replication-bacteriophage lambda." *Curr. Top. Microbiol. Immunol.* 78:201-237.
- Slocombe, P., A. Easton, P. Boseley, and D. C. Burke. (1982). "High level expression of an interferon $\alpha 2$ gene cloned in phage M13mp7 and subsequent purification with a monoclonal antibody." *Proc. Natl. Acad. Sci. USA*. 79:5455-5459.
- Smith, K. M. (1956). "The structure of insect virus particles." *J. Biophys. Biochem. Cytol.* 2:301-306.
- Smith, K. M. (1958). "A study of the early stages of infection with the *Tipula* iridescent virus." *Parasitology*. 48:459-462.
- Smith, K. M., and C. F. Rivers. (1959). "Cross-inoculation studies with the *Tipula* iridescent virus." *Virology*. 9:140-141.
- Smith, K. M., and R. C. Williams. (1958). "Insect viruses and their structure." *Endeavour*. 17:12-21.
- Smith, K. M., G. J. Hills, and C. F. Rivers. (1961). "Studies on the cross-inoculation of the *Tipula* iridescent virus." *Virology*. 13:233-241.

- Stoltz, D. B. (1973). "The structure of icosahedral cytoplasmic deoxyriboviruses II. An alternative model." *J. Ultrastruct. Res.* **43**:58-74.
- Sutcliffe, J. G. (1978). "pBR322 restriction map marked from the DNA sequence. Accurate DNA size markers up to 4361 nucleotide pairs long." *Nucleic Acids Res.* **5**:2721-2728.
- Sutcliffe, J. G. (1979). "Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322." *Cold Spring Harbor Symp. Quant. Biol.* **43**:77-90.
- Tabor, S., and C. C. Richardson. (1985). "A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes." *Proc. Natl. Acad. Sci. USA.* **82**:1074-1078.
- Tajbakhsh, S. (1988). "Isolation and characterization of the *Tipula* iridescent virus capsid gene." Ph.D Thesis, Biology Department, Carleton University, Ottawa, Canada.
- Tajbakhsh, S., and V. L. Seligy. (1989). "Molecular biology of *Tipula* iridescent virus." in *Molecular Biology of Iridoviruses*, edited by G. Darai.
- Tajbakhsh, S., M. J. Dove, P. E. Lee, and V. L. Seligy. (1986). "DNA components of *Tipula* iridescent virus." *Biochem. Cell Biol.* **64**:495-503.
- Tal, M., A. Silverstein, and E. Nusser. (1985). "Why does Coomassie Brilliant Blue R interact differently with different proteins? A practical answer." *J. Biol. Chem.* **260**:9976-9980.
- Thomas, R. S. (1961). "The chemical composition and particle weight of *Tipula* iridescent virus." *Virology.* **14**:240-252.
- Thomas, R. S., and R. C. Williams. (1961). "Localization of DNA and protein in *Tipula* iridescent virus (TIV) by enzymatic digestion and electron microscopy." *J. Biophys. Biochem. Cytol.* **11**:15-29.
- Tobin, H., T. Staehelin, and J. Gordon. (1979). "Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications." *Proc. Natl. Acad. Sci. USA.* **76**:4350-4354.
- Van Etten, J. L., D. E. Burbank, D. Kuczarski, and R. H. Meints. (1982). "Virus infection of culturable *Chlorella*-like algae and development of a plaque assay." *Science.* **219**:994-996.
- Venkatesan, S., A. Gershowitz, and B. Moss. (1982). "Complete nucleotide sequences of two adjacent early vaccinia virus genes located within the inverted terminal repetition." *J. Virol.* **44**:637-646.
- Venkatesan, S., B. M. Baroudy, and B. Moss. (1981). "Distinctive nucleotide sequences adjacent to multiple initiation and termination sites of an early vaccinia virus gene." *Cell.* **125**:805-813.
- Vinuela, E. (1985). "African swine fever virus." *Curr. Top. Microbiol. Immunol.* **116**:151-170.
- Wagner, G. W., J. D. Paschke, W. R. Campbell, and S. R. Webb. (1974). "Proteins of two strains of mosquito iridescent virus." *Intervirology.* **3**:97-105.

- Weinstock, G. M., C. apRhys, M. L. Berman, B. Hampar, D. Jackson, T. J. Silhavy, J. Weisemann, and M. Zweig. (1983). "Open reading frame expression vectors: A general method for antigen production in *Escherichia coli* using protein fusions to β -galactosidase." *Proc. Natl. Acad. Sci. USA*. 80:4432-4436.
- Weir, J. P., and B. Moss. (1983). "Nucleotide sequence of the vaccinia virus thymidine kinase gene and the nature of spontaneous frameshift mutations." *J. Virol.* 44:530-537.
- Weir, J. P., and B. Moss. (1987a). "Determination of the transcriptional regulatory region of a vaccinia virus late gene." *J. Virol.* 61:75-80.
- Weir, J. P., and B. Moss. (1987b). "Determination of the promoter region of an early vaccinia virus gene encoding thymidine kinase." *Virology*. 158:206-210.
- Williams, R. C., and K. M. Smith. (1957). "A crystallizable insect virus." *Nature*. 179:119-120.
- Williams, R. C., and K. M. Smith. (1958). "The polyhedral form of the *Tipula* iridescent virus." *Biochem. Biophys. Acta*. 28:464-469.
- Willis, D. B. (1985). "Iridoviridae." *Curr. Top. Microbiol. Immunol.* 116:1-173.
- Willis, D. B. (1987). "DNA sequences required for *trans*-activation of an immediate-early frog virus 3 gene." *Virology*. 161:1-7.
- Willis, D. B., and A. Granoff. (1976). "Macromolecular synthesis in cells infected by frog virus 3. V. The absence of polyadenylic acid in the majority of virus specific RNA species." *Virology*. 73:543-547.
- Willis, D. B., and A. Granoff. (1980). "Frog virus 3 DNA is heavily methylated at CpG sequences." *Virology*. 107:250-257.
- Willis, D. B., D. Foglesong, and A. Granoff. (1984a). "Nucleotide sequence of an immediate-early frog virus 3 gene." *J. Virol.* 53:905-912.
- Willis, D. B., R. Goorha, and A. Granoff. (1984b). "DNA methyltransferase induced by frog virus 3." *J. Gen. Virol.* 49:86-91.
- Willis, D. B., R. Goorha, and V. G. Chinchar. (1985). "Macromolecular synthesis in cells infected by frog virus 3." *Curr. Top. Microbiol. Immunol.* 116:77-106.
- Willis, D. B., R. Goorha, M. Miles, and A. Granoff. (1977). "Macromolecular synthesis in cells infected by frog virus 3. VII. Transcriptional and post-transcriptional regulation of virus gene expression." *J. Gen. Virol.* 24:326-342.
- Wittek, R. (1982). "Organization and expression of the poxvirus genome." *Experientia*. 38:285-297.
- Wood, H. A. (1980). "*Autographa californica* nuclear polyhedrosis virus-induced proteins in tissue culture." *Virology*. 102:21-27.
- Wrigley, N. G. (1969). "An electron microscope study of the structure of *Sericesthis* iridescent virus." *J. Gen. Virol.* 5:123-134.

- Wrigley, N. G. (1970). "An electron microscope study of the structure of *Tipula* iridescent virus." *J. Gen. Virol.* 6:169-173.
- Xeros, N. (1954). "A second virus disease of the leatherjacket, *Tipula paludosa*." *Nature.* 174:562-563.
- Xeros, N. (1964). "Development of the *Tipula* iridescent virus (TIV)." *J. Insect Path.* 6:261-283.
- Yanisch-Perron, C. J., J. Vieira, and J. Messing. (1985). "Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC 19 vectors." *Gene.* 33:103-119.
- Young, R. A., and R. W. Davis. (1983a). "Efficient isolation of genes using antibody probes." *Proc. Natl. Acad. Sci. USA.* 88:1194-1198.
- Young, R. A., and R. W. Davis. (1983b). "Yeast RNA polymerase II genes: isolation with antibody probes." *Science.* 222:778-782.
- Younghusband, H. B., and P. E. Lee. (1969). "Virus-cell studies of *Tipula* iridescent virus in *Galleria mellonella* (L.) I. Electron microscopy of infection and synthesis of *Tipula* iridescent virus in hemocytes." *Virology.* 38:247-254.
- Yuan, R. (1981). "Structure and mechanism of multi-functional restriction endonucleases." *Ann. Rev. Biochem.* 50:280-315.
- Yuen, L., and B. Moss. (1987). "Oligonucleotide sequence signaling transcriptional termination of vaccinia virus early genes." *Proc. Natl. Acad. Sci. USA.* 84:6417-6421.
- Yule, B. G., and P. E. Lee. (1973). "A cytological and immunological study of *Tipula* iridescent virus-infected *Galleria mellonella* larval hemocytes." *Virology.* 51:409-423.
- Zabin, I., and A. V. Fowler. (1978). "β-galactosidase, the lactose permease protein, and thiogalactoside transacetylase." in *The Operon*, edited by J. H. Miller and W. S. Reznikoff. Cold Spring Harbor Laboratory, pp. 89-123.

APPENDIX A. Electrophoresis buffers, media and antibiotics.**ELECTROPHORESIS BUFFERS****TAE (50X)**

Per liter:

Tris	242	g
Glacial acetic acid	57.1	ml
0.5 M Na ₂ EDTA pH 8	100	ml

TBE (10X)

Per liter:

Tris	108	g
Boric acid	55	g
Na ₂ EDTA	9.3	g

LIQUID MEDIA**B Medium**

Per liter:

Bactotryptone (Difco)	10	g
NaCl	8	g
Sterilized by autoclaving.		

L Medium

Per liter

Bactotryptone (Difco)	10	g
Yeast Extract (Difco)	5	g
NaCl	0.5	g
Glucose (BDH)	2	g
1M NaOH	4	ml

Sterilized by autoclaving.

NZY Medium

Per liter

Yeast Extract (Difco)	5	g
Casamino acids (Difco)	10	g
NaCl	5	g
MgCl ₂ ·6H ₂ O	2	g

Sterilized by autoclaving.

SOB Medium

Per liter

Bactotryptone (Difco)	20	g
Yeast Extract (Difco)	5	g
NaCl	0.584	g
KCl	0.186	g

Sterilize by autoclaving. Once cool the medium is completed by adding 10 ml each of filter sterilized 1M MgCl₂ and 1M MgSO₄.

SOC Medium

Identical to SOB medium except that 1/50th volume of filter sterilized 1M glucose is added.

2X YT Medium

Per liter

Bactotryptone (difco)	16	g
Yeast Extract (Difco)	10	g
NaCl	5	g

Sterilize by autoclaving.

MEDIA CONTAINING AGAROSE**Top Agarose**

Top agarose was prepared by adding 5 g of SeaKem ME agarose to one liter of B medium, and is sterilized by autoclaving.

NZY Top Agarose

NZY top agarose was prepared by adding 7 g of SeaKem ME agarose to one liter of NZY medium, and is sterilized by autoclaving.

BLUE/WHITE COLOUR ASSAY SOLUTIONS**IPTG**

Prepared a 100 mM solution by dissolving 23.8 mg solid Isopropylthio- β -D-galactoside in 1 ml of water. The resulting solution was filter sterilized and stored in the dark at 4°C.

BLUO-GAL

Prepared a 2% solution by dissolving 20 mg halogenated indolyl- β -D-galactoside in 1 ml dimethylformamide. The resulting solution was stored in the dark at 4°C.

ANTIBIOTICS**Ampicillin**

A 100X stock solution was prepared by dissolving 50 mg solid ampicillin (Sigma) in 10 ml water. The resulting solution was filter sterilized and stored at 4°C.

Kanamycin

A 100X stock solution was prepared by dissolving 50 mg solid Kanamycin Sulphate (Sigma) in 10 ml of water. The resulting solution was filter sterilized and stored at 4°C.

APPENDIX B. Preparation of polyacrylamide protein gels (Laemmli, 1970).

SOLUTION	10% BASE	4% STACK
acrylamide stock solution ^a	10.5ml	2.6ml
1 M Tris·Cl pH 8.8	11.9ml	---
1 M Tris·Cl pH 6.9	---	2.5ml
H ₂ O	8.75ml	14.7ml
20% SDS	156ul	100ul
TEMED	47ul	30ul
10% Ammonium Persulfate	125ul	125ul

^a Acrylamide stock solution was prepared by dissolving 29.2 gm acrylamide and 0.8 gm N',N',-Bis-methylene-acrylamide in 100 ml water, filtering, and was stored at 4°C in the dark.

APPENDIX C. Location of ORFs A to N on the region of the TIV genome sequenced in this thesis.

STRAND	ORF	START(nt)	END(nt)
L96 RNA sense	A	266	472
	B	379	603
	C	1361	1549
	D	1943	2416
	E	3365	3814
	F	4070	4254 (incomplete)
L96 cDNA sense	G	666	848
	H	1007	1321
	I	1388	1588
	J	1721	2392
	K	2474	2782
	L	3390	3617
	M	3809	4030
	N	4018	4254 (incomplete)

END

1 8-02-90

FIN