Analysis of the Functions and Products of the Yeast Retrotransposon Ty



by

Michael H. Malim

Thesis submitted for the degree of

Doctor of Philosophy

at the

University of Oxford

.

St. Catherine's College

September, 1987

ANALYSIS OF THE FUNCTIONS AND PRODUCTS OF THE YEAST RETROTRANSPOSON TY.

Michael H. Malim, St. Catherine's College, Oxford.
Thesis for Doctor of Philosophy. September 1987.

ABSTRACT

There are 30-35 copies of the retrotransposon Ty in the haploid genome of most laboratory strains of Saccharomyces cerevisiae. Ty elements are 5.9 kb long, have LTRs of 340 bp and produce a genomic RNA that is 5.7 kb in length, through which it transposes. They contain two ORFs; the TYA gene of the class I element Ty1-15 is 1,319 nucleotides long whilst its TYB gene, which possesses the consensus sequences of retroviral acid proteases, reverse transcriptases and integrases, is 3,984 nucleotides long. TYB overlaps TYA by 38 nucleotides at its 5' end and is expressed as a TYA:TYB fusion protein.

In an attempt to asign some in vivo functions to the Ty encoded proteins, the whole transcriptional unit of Ty1-15 was overexpressed in yeast. Electron micrographs of the overexpressing cells revealed an abundance of cytoplasmic, ~60 nm virus-like particles (Ty-VLPs). These contain the 5.7 kb Ty RNA and have an associated reverse transcriptase activity. Their major protein is p2 (48 kD), which is a proteolytic derivative of the primary translation product of TYA, p1 (50 kD), a protein that is also assembled into particles.

A truncated p1, encoded by a TYA gene shortened by 60 codons at its 3' end, still aggregates into particles. Furthermore, the addition of α2-IFN (20 kD) onto its carboxy terminus, to produce a p1-IFN fusion protein of ~68 kD, does not interfere with its particle forming properties. The hybrid VLPs are easy to purify and elicit an antibody response in rabbits to the Ty and IFN epitopes.

A series of fragments spanning Ty1-15 were inserted into the packaging analysis vector, pMA924. The association of the resulting PGK-Ty-IFN hybrid RNAs with Ty-VLPs, in the presence of abundant Ty encoded proteins, was analysed by Northern blotting and used to assess the packaging capabilities of Ty RNA. Unexpectedly, RNA sequences that direct specific packaging are produced from regions throughout Ty1-15. The signals corresponding to TYA are probably located within two regions of the genomic RNA, 5' to nucleotide 236 and 3' to nucleotide 737; those corresponding to TYB are in at least two, currently undefined, areas.

The TYB gene of the class II element Ty1-17 overlaps TYA by 44 nucleotides. The fusion of α2-IFN cDNA fragments into all three reading phases of TYB followed by a Western blot analysis demonstrated that, like the TYB gene of Ty1-15, it too is expressed as a TYA:TYB fusion protein.

Thesis Supervisor: Dr. S. M. Kingsman.

ACKNOWLEDGEMENTS

I would like to acknowledge the following, for aiding and abetting me during my three years at Oxford. They have all contributed towards making this thesis, and the research that is described within it, appear in this final form.

I am indebted to my Mum and Dad for supporting me throughout my life as a student, both in Oxford and, as an undergraduate, in Bristol. I hope that I can repay their generosity in the years to come.

I am grateful to my supervisors Alan and Sue Kingsman for giving me a place in their laboratory and for directing my research. I also thank Alan Kingsman for sorting out the irregularities in this thesis prior to its submission and for his encouragement and advice with respect to my future career in science. I thank the SERC for funding me whilst at Oxford and for financing my visit to the Thirteenth International Conference on Yeast Genetics and Molecular Biology (Banff, Alberta, 1986), for which I also thank St. Catherine's College and the Graduate Studies Committee for their support.

I am very grateful to Clive Stanway for running the first draught of this thesis past himself and for his many helpful comments and suggestions; to Rebecca Oakey for her continual encouragement during the production of this thesis and comments on punctuation! to Jackie Hewitt for the typing and to Frank Caddick for the photography and the artwork in Figure 7.8.

From the Kingsman group itself, I thank Jane Mellor and Sally Adams for their collaborations and assistance with Chapters 3 and 7, respectively. However, everyone played a part and helped to maintain my sanity at some of the more trying moments, though special mention should go to: Steve Harrison for his Yorkshire wit; Peter Rathjen for assorted "testing"; Clive Stanway, Jill Ogden and Sandra Fulton for helpful discussions and some of the Ty plasmids (S.F.); and Donald Tipper for keeping things in perspective!

From outside of the lab, the assistance of Bob Sim for immunising the rabbits, Keith Gull for the electron microscopy and Ian Jones for providing the purified influenza viruses was gratefully received.

TABLE OF CONTENTS

		Page No.
Chapter 1.	**Introduction**	
1.1	General Introduction: Eukaryotic Genome Organisation	1
1.2	Highly Repeated Sequences	1
1.3	Middle Repeated Sequences	2
1.3.1	Retroposons	4
1.3.1.1	Processed Pseudogenes	4
1.3.1.2	Small Nuclear RNA Pseudogenes	5
1.3.1.3	Short Interspersed Repetitive DNA Sequences (SINEs)	5
1.3.1.4	Long Interspersed Repetitive DNA Sequences (LINEs)	6
1.3.2	Transposons without Long Terminal Repeats	7
1.3.2.1	Transposable Elements with Long Inverted Repeats	7
1.3.2.2	Transposable Elements with Short Inverted Terminal Repeats	8
1.3.3	Transposons with Long Terminal Repeats and their Relationship to Retroviral Proviruses	10
1.3.3.1	Retroviruses	11
1.3.3.2	Copia-Like Elements of *Drosophila*	19
1.3.3.3	TED Elements of *Trichopulsia ni*	22
1.3.3.4	Intracisternal A Particles of Rodents	23
1.4	The Yeast Retrotransposon Ty	24
1.4.1	General Features	25
1.4.2	Recombination between Ty and Delta Sequences	27
1.4.3	Transcription	28
1.4.4	Genetic Effects	30
1.4.5	Ty Proteins and the Strategy of Over-expression	32
1.4.6	Transposition	35
1.5	Summary and Thesis Objectives	36
Chapter 2.	**Materials and Methods**	
2.1	*E. coli* Strains	39
2.2	Yeast Strains	39
2.3	*E. coli* Media	39
2.4	Yeast Media	40

2.5	Plasmid DNA Isolation from *E. coli*	41
2.5.1	Rapid Procedure ("miniprep")	41
2.5.2	Large Scale Plasmid Preparation: Triton Method	42
2.5.3	The Direct Transfer of Plasmid DNA onto Nitrocellulose Filters	44
2.6	Plasmid DNA Preparation from Yeast	45
2.6.1	Rapid Isolation of Plasmid DNA from Yeast	45
2.6.2	Large Scale Preparation of Plasmid DNA from Yeast	46
2.7	Gel Electrophoresis of DNA	48
2.7.1	Agarose Gel Electrophoresis	48
2.7.2	Polyacrylamide Gel Electrophoresis	49
2.8	Recovery of DNA Fragments from Gels	5Ø
2.8.1	DNA Fragments from Agarose Gels	5Ø
2.8.1.1	Using Glycerol	51
2.8.1.2	Using Dialysis Tubing	52
2.8.2	DNA Fragments from Polyacrylamide Gels	52
2.9	DNA Manipulations	53
2.9.1	Restriction Digestion	53
2.9.2	Dephosphorylation	55
2.9.3	Fill-in Reaction	55
2.9.4	Ligation of Sticky Ends	56
2.9.5	Ligation of Blunt Ends	57
2.9.6	Bal31 Digestion	57
2.1Ø	Determination of Nucleotide Sequences	58
2.11	*E. coli* Transformation	59
2.11.1	AKEC28	59
2.11.2	HB1Ø1	6Ø
2.12	Yeast Transformation	6Ø
2.13	Preparation of Proteins from Yeast	62
2.14	Preparation of Virus-Like Particles (VLPs) from Yeast	63
2.15	Assay for Reverse Transcriptase Activity	65
2.16	Western Blot Analysis	66
2.16.1	SDS-Polyacrylamide Gel Electrophoresis	66
2.16.2	Coomassie Blue Staining	68
2.16.3	Transfer of Proteins to Nitrocellulose	68
2.16.4	Hybridisation with Double Stranded DNA	69
2.16.5	Hybridisation with ^{125}I Labelled Monoclonal Antibody	7Ø

2.16.6 Hybridisation with Non-Radiolabelled Antibody: Detection with an Immuno-enzymatic Reaction 70
2.16.7 Immunisation of Rabbits with VLPs 72
2.16.8 Preparation of Neat Serum from Blood 72
2.16.9 Purification of IgG 72

2.17 Electron Microscopy 73

2.18 Northern Blot Analysis 74
2.18.1 Preparation of Total Yeast RNA 74
2.18.2 Preparation of RNA from VLPs 75
2.18.3 Running the Gel 76
2.18.4 Transfer of the RNA from the Gel to Nitrocellulose 78
2.18.5 "Nick Translation" 79
2.18.6 Hybridisation 80
2.18.7 Quantitation 81

Chapter 3. The Identification and Characterisation of Ty Virus-Like Particles in Yeast

3.1 Introduction 82

3.2 The Identification of Ty Virus-Like Particles in Yeast 82

3.3 The Composition of Ty Virus-Like Particles 84

3.4 The Properties of the Ty-VLP Associated Reverse Transcriptase 87

3.5 Discussion and Subsequent Results 89

Chapter 4. RNA Homologous to the TYA Region of Tyl-15 is Packaged into Ty Virus-Like Particles: the Development of an Assay System

4.1 Introduction: Retroviral RNA Packaging and its Exploitation 94

4.2 Experimental Approach to the Analysis of Ty RNA Packaging Signals 97

4.3 RNA from the 5' end of Tyl-15 is Packaged into Ty-VLPs 99

4.4 The Evaluation of Packaging Directed by the 11 Fragment 102

4.5 The Construction of a 3' to 5' Deletion Series through the 11 Fragment 106

4.6 The Packaging Analysis of the 11 Fragment Deletion Series 108

4.7 Discussion 110

Chapter 5. Sequences that are Capable of Directing RNA Packaging are Dispersed Throughout Ty1-15

5.1 Introduction 113

5.2 Plasmid Constructions 114

5.3 Packaging Signals Reside Within RNAs Derived from Dispersed Regions of Ty1-15 115

5.4 Conclusions and Discussion 117

Chapter 6. Truncated p1 Proteins and p1-Interferon Fusion Proteins can both Assemble into Virus-Like Particles

6.1 Introduction 120

6.2 Plasmid Constructions 120

6.3 The Ability of Truncated p1 Proteins to Form Particles 122

6.4 The Ability of p1-IFN Fusion Proteins to Form Particles 124

6.5 Electron Microscopy and Discussion of Particle Formation 125

Chapter 7. Hybrid Ty Virus-Like Particles can Function as Self-Assembling Epitope Carriers

7.1 Introduction: Recombinant DNA Approaches Vaccine Development 128

7.2 The TYA:IFN Fusion Particles have Ty and Interferon Epitopes 132

7.3 The Ty:IFN Hybrid VLPs Elicit an Antibody Response Against Interferon 133

7.4 Hybrid VLPs Carrying a Viral Epitope 136

7.5 Conclusions and Discussion 137

Chapter 8. A Search for the cDNA Intermediates of Transposition

8.1 Introduction 139

8.2 Plasmid Constructions 143
8.3 The Isolation and Restriction Analysis of Plasmids that Confer Ampicillin Resistance 144
8.4 Discussion 147

Chapter 9. The TYB Gene of the Class II Element Tyl-17 is Expressed as a TYA:TYB Fusion Protein

9.1 Introduction 149
9.2 Plasmid Constructions 150
9.3 Western Blot Analysis of the TYB:IFN Fusions 151
9.4 Discussion 152

Chapter 10. Discussion

10.1 General Discussion and Evolutionary Considerations 154
10.2 Ty Elements as Selfish DNA 161
10.3 Future Investigations 163

References 166

ANALYSIS OF THE FUNCTIONS AND PRODUCTS OF THE YEAST RETROTRANSPOSON TY

CHAPTER 1. INTRODUCTION

1.1 General Introduction: Eukaryotic Genome Organisation

The work presented in this thesis concerns the Saccharomyces cerevisiae (or budding yeast) retrotransposon Ty. Ty is a middle repetitive DNA element and there are 30-35 copies per haploid genome.

The genomes of eukaryotes exist as chromosomes that have three categories of sequence. These are the highly-, middle-, and non-repetitive (or unique) sequences (Lewin, 1983, pp281-291). These categories were identified by reassociation kinetics, whereby the reassociation rate of denatured DNA is measured. A highly repeated sequence will reanneal the fastest because its concentration will be the highest. The non-repetitive sequences are the slowest to reanneal and are usually the major component of eukaryotic genomes. A fourth genomic component that is known as zero-time binding DNA is often present. These sequences reanneal instantaneously after denaturation because they are inverted repeats or palindromes; they follow rapid first order renaturation kinetics because these complementary sequences are on the same molecule.

1.2 Highly Repeated Sequences

Highly repetitive DNA is composed of short sequences that are repeated many times in tandem arrays. Consequently, this component of the genome is frequently assembled into extended regions of DNA that can exceed 100 kb in length (reviewed by Brutlag, 1980). The base composition of these repeated sequences

may be somewhat different from the rest of the genome. This characteristic allows it to be physically separated from the main band DNA by centrifugation through a CsCl density gradient: such separated material is called satellite DNA. In higher eukaryotes the satellite DNA tends to be located at the telomeres and centromeres of chromosomes. However, no function has been assigned to the satellites as yet (Lewin, 1983, pp378-391).

S. cerevisiae has very little satellite DNA. Yeast centromeres have been cloned and, in contrast to higher eukaryotes, they are not "buried" in a mass of satellite DNA. Furthermore, yeast centromeres have been sequenced and have been shown to possess limited homology to the higher eukaryotic satellites (Fitzgerald-Hayes et al., 1982). Yeast telomeres have been shown to contain simple repeats for a few hundred nucleotides (Walmsley et al., 1984; Shampay et al., 1984).

1.3 Middle Repeated Sequences

This component of the eukaryotic genome is quite different from the highly repetitive sequences. The essence of middle repetitive sequences is that they are interspersed throughout the genome, rather than existing in tandem arrays or clusters. The abundance of different classes of middle repetitive sequence can vary enormously, for example, there are nearly a million Alu repeats in the human genome (Hwu et al., 1986) but only 30-35 Ty elements in S. cerevisiae (Cameron et al., 1979). Any sequence that is represented more than once is, by definition, a

repetitive sequence. Therefore, repeated genes such as the 5S RNA genes of Xenopus and the mammalian globin gene family are repetitive DNA.

One process that is thought to have contributed towards the generation and accumulation of a certain amount of the repetitive DNA of eukaryotes is reverse transcription. The enzyme reverse transcriptase is unique in its ability to produce a DNA copy from an RNA template (Temin and Mizutoni, 1970; Baltimore, 1970). It is therefore conceivable that any cellular RNA molecule is capable of giving rise to a cDNA. Furthermore, the insertion of such a cDNA into the genome would render the original sequence repetitive. Indeed, it has been suggested by Baltimore (1985) that as much as ten percent of some eukaryotic genomes may have been produced by reverse transcription.

The eukaryotic middle-repetitive sequences that are dispersed throughout their respective genomes include elements that are known as retroposons and transposons; the events that result in their dispersal are called retroposition and transposition, the hall-mark of which is the retention of the element at its original genomic location. For simplification, eukaryotic transposons can be divided into two broad classes, those with Long Terminal Repeats (LTRs) and those without. The widespread distribution of these repetitive elements constitutes a clear distinction between them and clustered repeated sequences, for example, the mammalian globin gene family.

1.3.1 Retroposons

As the name suggests, the propagation of these repetitive elements has probably involved reverse transcription (for review, see Rogers, 1985) though mechanisms such as gene duplication, transposition and chromosomal rearrangements have almost certainly played a role. Retroposons have some general features; their 5' and 3' ends correspond to the termini of an RNA species, there is a repetitive (A)-rich tail at the 3' end and there are short direct terminal repeats of 8-19 bp of the flanking sequence at both the 5' and 3' ends. The reverse transcriptase is generally not coded for by the retroposon. A mechanism of self-priming has been suggested for the putative reverse transcription (Figure 1.1, Jagadeeswaran *et al.*, 1981). No insertion event of a retroposon has been observed *in vivo*. There are four basic categories of retroposon.

1.3.1.1 Processed Pseudogenes

A pseudogene is a DNA sequence that is originally derived from a functional gene. Once created, a non-functional pseudogene may accumulate numerous mutations. Most of the pseudogenes of the mammalian globin gene family were created by gene duplication and therefore, still contain their introns (reviewed by Little *et al.*, 1981). The mouse α3 pseudogene, however, is at a different genetic locus (chromosome 15 rather than chromosome 11) and has lost its introns (Leder *et al.*, 1981): such sequences are called processed pseudogenes. The

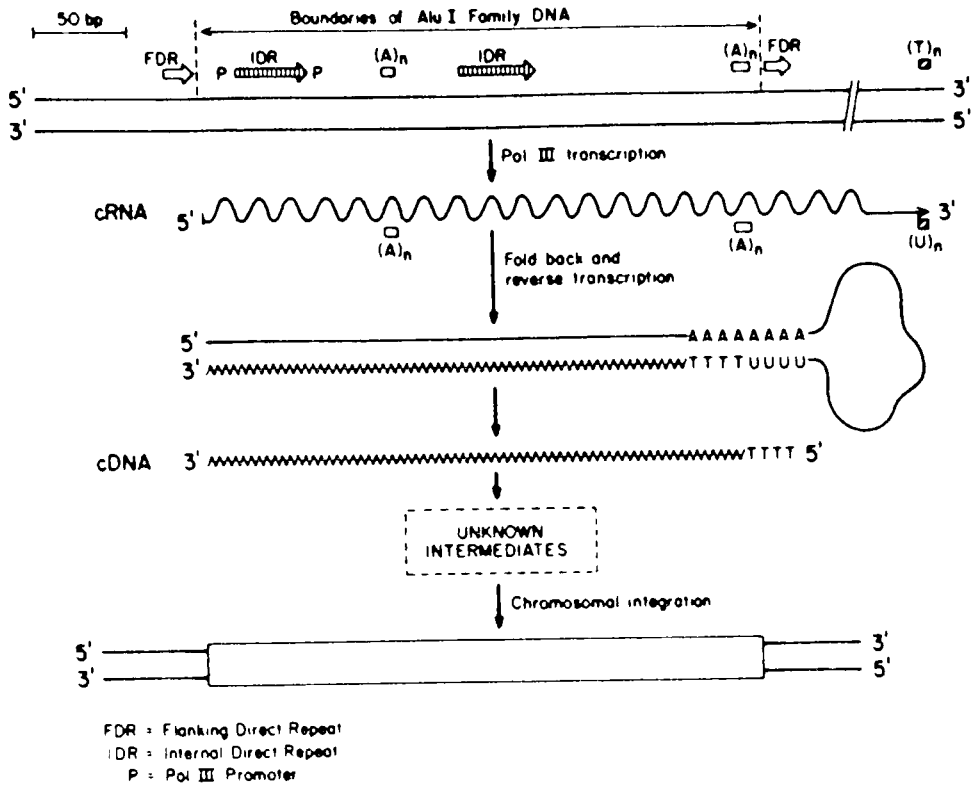


Figure 1.1. Schematic diagram illustrating the structure, repeated sequences and proposed mechanism of retroposition of an Alu DNA sequence (taken from Jagadeeswaran et al., 1981).

dispersal of these sequences must involve RNA because it is only RNA that is spliced, therefore reverse transcriptase is implicated (Sharp, 1983). Processed pseudogenes are normally found in mammalian genomes, an exception being the calmodulin pseudogene of chickens (reviewed by Wagner, 1986). The reverse transcriptase that is necessary for the generation of processed pseudogenes is probably provided by the endogenous retroviruses (Section 1.3.3.1) that are found in the genomes of mammals and birds. Processed pseudogenes are generally non-functional, however McCarey and Thomas (1987) have shown that the autosomal PGK-2 gene of humans has the characteristics of a retroposon and is expressed during late spermatogenesis when the X-linked PGK-1 is temporally inactivated.

1.3.1.2 Small Nuclear RNA Pseudogenes

The snRNA species U1 to U6, each hybridise to 10^2-10^3 sequences in mammalian genomes. Most of these repeats are pseudogenes, many having the characteristics of retroposons. Further sequence data have demonstrated that many of these pseudogenes possess variable truncated 3' ends, indicating various modes of propagation (reviewed by Rogers, 1985).

1.3.1.3 Short Interspersed Repetitive DNA Sequences (SINEs)

The prototypes for these mammalian repetitive sequences are the AluI repeats of primates and the B1 repeats of rodents. The

single copies of Alu DNA are approximately 300 bp in length and have a diagnostic AluI restriction site: B1 elements are approximately 130 bp in length (reviewed by Jagadeeswaran et al., 1981; Rogers, 1985). These elements both have the classical retroposon characteristics and are presumed to have spread via an RNA intermediate (Figure 1.1). The abundance of these repeats is of the order of 10^5 per haploid genome (Rogers, 1985; Hwu et al., 1986). They are templates for RNA polymerase III in vitro, but are not expressed as discrete units in Hela cells (Paulson and Schmid, 1986) even though they are present in the high molecular weight heterogeneous nuclear RNA. The origin of the AluI family is believed but not proven to be the 7SL RNA gene (Ullu and Tschudi, 1984). No obvious role other than self preservation has been assigned to SINEs, but a possible role in genome rearrangement has been supported by the identification of an Alu-Alu mediated recombination event that resulted in a seven exon duplication in the LDL receptor gene of a patient with Famial Hypercholesterolaemia (Lehrman et al., 1987).

1.3.1.4 Long Interspersed Repetitive DNA Sequences (LINEs)

The prototypes for these mammalian repetitive sequences are the L1 family of primates and the L1Md family of mouse (reviewed by Rogers, 1985). In contrast to SINEs these elements are 6-7 kb in length and have a typical abundance of 10^4-10^5 copies per haploid genome (Hwu et al., 1986). LINEs have the usual structural characteristics of retroposons but unlike SINEs, are

transcribed by RNA polymerase II (Skowronski and Singer, 1985). Sequencing of LINEs has revealed some interesting features (for L1Md, refer to Loeb et al., 1986; and for L1, refer to Fanning and Singer, 1987). For instance, there are homologies to the consensus sequences of retroviral reverse transcriptases (Toh et al., 1983) and nucleic acid binding proteins (Berg, 1986) as well as the transcription factor TFIIIA (Miller et al., 1985). These observations have prompted Fanning and Singer (1987) to propose a common evolutionary ancestor for retroviruses, retrotransposons (see later) and LINEs. LINEs are therefore distinct from other retroposons as they code for a putative reverse transcriptase, this may or may not have a role in propagation. It now appears that LINEs are not restricted to mammals since the analysis of a Drosophila melanogaster I factor has demonstrated a similar sequence organisation (Fawcett et al., 1986).

1.3.2 Transposons without Long Terminal Repeats

The transposons that lack LTRs have very diverse structures, do not encode a reverse transcriptase and are therefore unlikely to rely upon an RNA intermediate for their dispersal (transposition).

1.3.2.1 Transposable Elements with Long Inverted Repeats

These elements have been studied the most in D. melanogaster. One class of elements are known as Fold-Back (FB) elements, some of which have been sequenced, but no regular

structure, apart from the inverted terminal repeats, has emerged. FB elements do transpose and may result in mutations of the white gene (Levis et al., 1982; Collins and Rubin, 1983). Similar FB-like elements have also been identified in the sea urchin (Liebermann et al., 1983).

Another type of element in D. melanogaster is the TE element. These transposons are the largest in eukaryotes and may be over 100 kb in length. TE elements are often flanked by FB elements: the DNA content of a given TE element (also known as composite transposons) may alter with time, as seen in certain TE elements close to the white gene (Paro, 1983).

1.3.2.2 Transposable Elements with Short Inverted Terminal Repeats

Some of the transposons in this class have been studied extensively and may have profound effects upon their host as a result of either transposition or genomic rearrangements mediated by the transposon.

The P elements of D. melanogaster have a size range of 500 bp up to the 2.9 kb of an autonomous element (O'Hare and Rubin, 1983). The full length element encodes a transposase activity that can trans-activate the smaller, truncated elements. Transposition of P elements is repressed in P-strain flies (contain P elements) but is allowed in M-strain flies (contain no P elements) (Engels, 1983). Thus, transposition only occurs when P elements are introduced into an M-cytotype egg; a syndrome known as P-M hybrid dysgenesis ensues, this is characterised by

high rates of sterility, mutation and chromosomal rearrangements (Kidwell *et al.*, 1977). In addition, transposition is restricted to germline tissue, hence it is tissue specific and genetically regulated. Laski *et al.* (1986) have shown that three splicing events are required for the expression of the transposase, the third of which is germ-line specific and requires a factor only present in M-strain flies. The 88 kD transposase has been purified and shown to catalyse P element excision (Rio *et al.*, 1986). P elements have also been used for germ-line transformation of *Drosophila* (Spradling and Rubin, 1982).

The first transposable elements were discovered in maize over thirty years ago by McClintock (1950). The best characterised family include the Dissociation (Ds) and the Activator (Ac) elements which can cause a variety of mutations when inserted into different genetic loci of the host genome. There is one parallel with the P elements of *Drosophila*: Ac elements can transpose freely, but Ds elements will only transpose in the presence of an Ac element. Sequencing has shown that Ac elements are 4,563 bp long and contain three divergent Open Reading Frames (ORFs). Ds elements differ from Ac elements by variable internal deletions, presumably deletions in the transposase coding region prevent Ds elements catalysing their own transposition (reviewed by Doring and Starlinger, 1984).

Other plant transposable elements include the Tam transposons of *Antirrhinum majus* (reviewed by Coen and Carpenter, 1986). These elements are variable in size and have been studied as mutators of the *nivea* and *pallida* genes: these genes code for two of the enzymes in the pathway for anthocyanin synthesis;

consequently, any mutation of them may affect flower pigmentation. A variety of stable and unstable mutations that may be caused by element excision or chromosomal rearrangements have been identified (refer to Coen et al., 1986 for Tam3 in pal^{rec}). No Tam element has been fully sequenced as yet, but Tam3 is transcribed and contains at least one ORF.

1.3.3 Transposons with Long Terminal Repeats and their Relationship to Retroviral Proviruses

The members of this class of repetitive elements have several distinguishing features that set them apart from the elements discussed so far, Figure 1.2 shows some of these. The elements are flanked by Long Terminal Repeats (LTRs) that have similar sequences in a given element. Each LTR often has its own Inverted Terminal Repeats (ITRs). The terminal bases are normally 5'-TG---CA-3'. There is a Direct Repeat (DR) of the flanking DNA at the insertion site, the length of which is characteristic of the element (for example, Ty elements are flanked by a 5 bp duplication of the insertion site DNA, Farabaugh and Fink, 1980). A sequence 3' to the 5' LTR is complementary to a cellular tRNA, which can act as a primer for reverse transcription (in contrast to the proposed self-priming of Alu element retroposition). These gross structural features are reminiscent of integrated retroviral proviruses (Varmus, 1983). Retroviruses code for their own reverse transcriptase and rely upon it for their horizontal transmission. Interestingly,

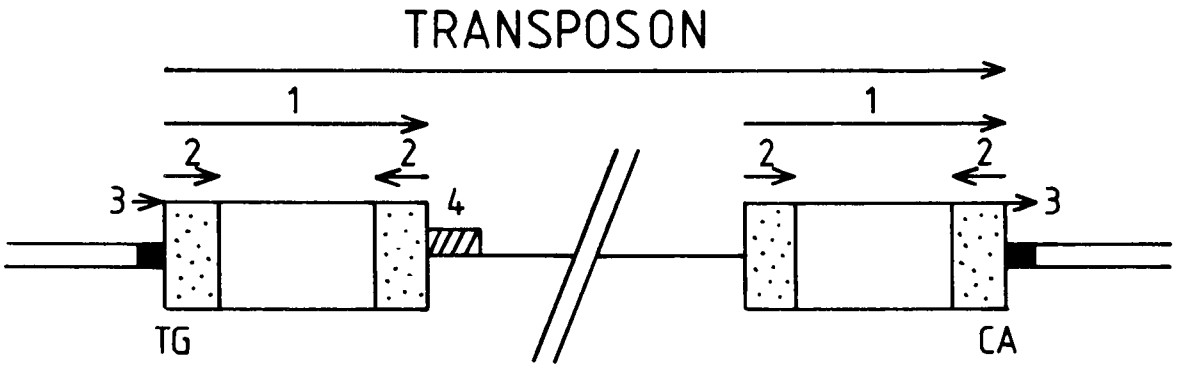


Figure 1.2. General features of a eukaryotic transposon with Long Terminal Repeats (LTRs). 1, LTRs; 2, Inverted Terminal Repeats (stippled boxes); 3, short duplications of the target site DNA (solid boxes); 4, proposed tRNA primer binding site (hatched box).

the sequencing of the eukaryotic transposons with LTRs, such as the copia-like elements of Drosophila, Ty of yeast and IAP genomes of rodents, revealed homologies to the retroviral reverse transcriptase consensus sequence (Mount and Rubin, 1985; Hauber et al., 1985; Clare and Farabaugh, 1985; Ono et al., 1985; Toh et al., 1983); consequently, the term retrotransposon is frequently used to describe this class of eukaryotic transposons. Further studies have revealed similarities in expression strategies and protein function between the integrated proviruses and retrotransposons.

1.3.3.1 Retroviruses

Retroviruses are a large class of animal viruses united by the replication of their RNA genomes through double stranded DNA intermediates (reviewed by Varmus, 1983). Reverse transcriptase is therefore essential for this life cycle. The genomes of both mammalian and avian cells contain repeated sequences that are known as endogenous retroviruses (or proviruses), these are similar to the integrated retroviral proviruses except that they are often defective in replication or expression (reviewed by Jaenisch, 1983). For simplicity, the two will be regarded as synonymous, though endogenous proviruses are inherited vertically as Mendelian genes whilst retroviruses are transmitted horizontally and vertically. Inspite of their defective nature, endogenous proviruses may still give rise to mutations that are normally associated with retroviral infections, for example, a spontaneous coat-colour mutation in mice has been shown to be

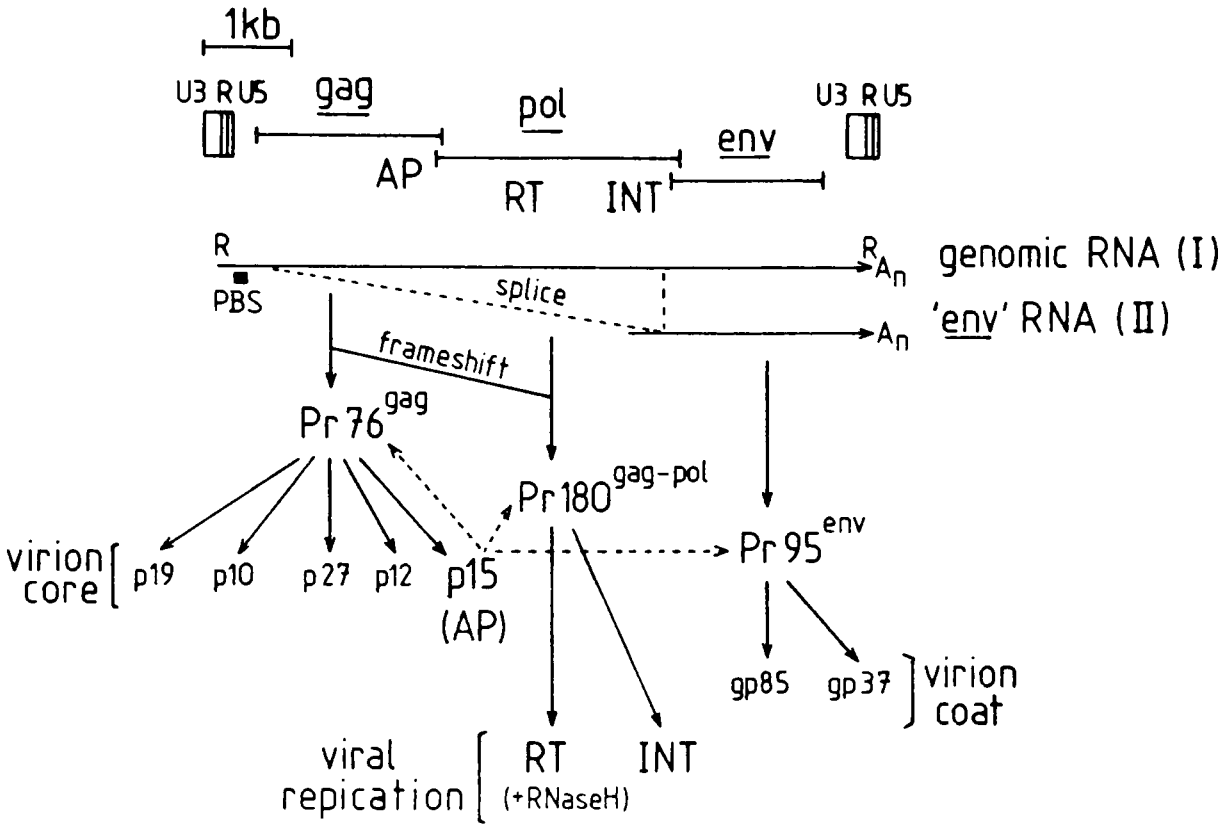


Figure 1.3. Genome organisation and expression strategy of the integrated provirus of Avian Leukosis Virus (ALV). AP, acid protease; RT, reverse transcriptase; INT, integrase; R, 50 nucleotide terminal redundancy of transcript I; Pr, polyprotein precursor; p, protein; gp, glycoprotein.

associated with the insertion of an endogenous provirus (Jenkins et al., 1981).

The genomic organisation and expression strategy of a typical retrovirus, in this case Avian Leukosis Virus (ALV), is shown in Figure 1.3. Many of the observations that are described here were made with Rous Sarcoma Virus (RSV). It is thought that ALV was the progenitor retrovirus of RSV, before it acquired the src gene by transduction (Schwartz et al., 1983), consequently such results relate to ALV. The LTRs are divided into three distinct regions known as U3, R and U5. The unique region between the LTRs of ALV contains three genes known as gag, pol and env. Figure 1.4 shows that retroviruses such as Human Immunodeficiency Virus (HIV) have additional genes, the functions of which are only now being elucidated (see later).

The genomic RNA of ALV extends from the U3-R junction of the 5' LTR to the R-U5 junction 3' LTR, therefore it has the R region present at both ends and is said to be terminally redundant (see later). The LTRs contain at least some of the consensus sequences that are important in eukaryotic RNA polymerase II transcription. There is a TATA box (Corden et al., 1980) at 20-30 bp upstream of the RNA start site, a CAAT box (Benoist et al., 1980) within 110 bp of the RNA start site, and an enhancer element (Khoury and Gruss, 1983) further upstream. These cis-acting transcription components are all contained in the U3 region and have been found to be sufficient for normal transcription from the RSV LTR (Cullen et al., 1985), however the enhancer is not as "flexible" as that of SV40 (Banerji et

al., 1981) in that its relative position to the other promoter components seems to be important if normal transcription is going to occur. Other components of promoters, such as binding sites for the transcription factor Sp1, have been identified in the LTRs of retroviruses such as HIV (Jones et al., 1986). LTRs also contain the signals responsible for transcription termination and polyadenylation (usually AATAAA, Proudfoot and Brownlee, 1974), though these are only active in the 3' LTR. Transcription proceeds from the 5' LTR, through the genome, and into the 3' LTR; this prevents the 3' LTR acting as a promoter, a phenomenon that is known as Transcriptional Overlap Interference (Cullen et al., 1984). Recently, cis-acting sequences in the gag region of RSV have been implicated in the induction of maximum transcription (Arrigo et al., 1987).

The host cell is generally able to exert some control over retroviral expression. For example, the transcriptional enhancer of Moloney Murine Sarcoma Virus is responsive to glucocorticoid hormones (Miksicek et al., 1986) and the transcription of endogenous proviruses is under developmental control (reviewed by Jaenisch, 1983). It will be described in the following sections that a hallmark of retrotransposons, like integrated retroviral proviruses and endogenous proviruses (the retro-elements), is the host's ability to affect the gene expression of the element by the interactions of gene products from unlinked loci; for example, Ty element transcription is regulated by the mating status of the yeast cell (Section 1.4.3).

The gag, pol and env coding regions are all present in the full length genomic RNA (transcript I in Figure 1.3). A major

splicing reaction occurs to remove the majority of the gag and all of the pol coding regions from I to yield transcript II. Therefore env is the only complete ORF within transcript II. Transcripts I and II are normally present at roughly equal levels in provirus expressing cells.

The expression of transcript I to produce $Pr76^{gag}$ and of transcript II to produce $Pr95^{env}$ (Figure 1.3) are simple translation events involving ribosome binding at the first suitable AUG of the transcript, followed by elongation and termination. $Pr76^{gag}$ is proteolytically cleaved by the retrovirus encoded acid protease (see later) to yield a variety of smaller proteins, some of which are phosphorylated (Figure 1.3): these form the structural basis of the virion core (for the precise functions of the gag proteins, refer to Schwartz et al., 1983). $Pr95^{env}$ is also cleaved, as well as glycosylated, to produce the glycoproteins that form the envelope of the mature virion. The pol gene of murine retroviruses (for example, Moloney Murine Leukaemia Virus, Figure 1.4) codes for proteins that are enzymatic rather than structural: these are an acid protease (involved in gag and env maturation), a reverse transcriptase (produces a DNA copy of the genomic RNA) and an integrase (double stranded DNA endonuclease). In avian retroviruses, however, the acid protease is at the 3' end of gag as opposed to the 5' end of pol (Figures 1.3 and 1.4). These enzymatic proteins are not required at such high levels as the structural proteins and the virus regulates this by expressing pol as a gag-pol fusion protein, whereby the fusion occurs in less than ten percent of

the gag expression events. The pol gene of RSV overlaps the gag gene by 58 nucleotides and it is in the -1 reading phase with respect to gag (Schwartz et al., 1983): the $Pr180^{gag-pol}$ fusion protein is expressed via a ribosomal frameshifting event that brings these two genes into the same reading phase (Jacks and Varmus, 1985). Murine Leukaemia Virus (MLV) produces the gag-pol fusion by suppression of the amber stop codon that separates the two genes (Yoshinaka et al., 1985). Once the fusion protein has been produced, it is processed by the autocatalytic acid protease to produce the acid protease, the reverse transcriptase and the integrase. The striking feature of this complex expression cascade is that a large amount of information is packed into a small (8-9 kb) genome, this is a characteristic of many classes of virus.

Recent work has identified regions that trans-activate the transcription or the translation of some retroviruses. Broome and Gilbert (1985) have shown that a 124 amino acid protein produced by a 282 nucleotide alternative splice in the p19-p10 region of gag in RSV can activate LTR directed transcription. More recently, several novel splicing events have been identified that result in the expression of the trans-activating genes tatIII and art/trs of HIV (Figure 1.4) (Rosen et al., 1985; Sodroski et al., 1986; Cullen, 1986; Feinberg et al., 1986; Muesing et al., 1987). Whether these proteins affect transcription, translation, RNA stability, RNA splicing or even RNA transport is a matter of some controversy.

Once the RNA and proteins have been produced by a provirus they are assembled into virions. The mature retroviral particles

have a size range of 80-120 nm (Varmus, 1983). The order in which the assembly events take place differs from one retrovirus to another. The primary gag product, $Pr65^{gag}$, of Murine Leukaemia Virus is modified at the N-terminal glycine by myristylation, and this allows the protein to be aligned at the plasma membrane (Rein et al., 1986). "Immature" virions that consist of some cellular phospholipid, the precursor proteins $Pr65^{gag}$ and $Pr180^{gag-pol}$, the precursor glycoprotein $gPr90^{env}$ and two molecules of the full length genomic RNA then bud off. Processing of the precursor proteins now takes place, this is mediated by the retrovirus encoded autocatalytic acid protease (Figure 1.4) (Yoshinaka et al., 1985). The non-infective "immature" virions alter their morphology during this process and become "mature" virions that are capable of infection (Katoh et al., 1985). The spliced "env"-mRNA is not incorporated into virions because an RNA sequence that is downstream from its 5' splice donor site is required for efficient packaging of the viral RNA; this sequence is therefore absent from the spliced RNA (Mann et al., 1983; Chapter 4). The assembly pathway for avian retroviruses is rather different as it seems that some cleavage of $Pr76^{gag}$ is necessary for, and is simultaneous with, virion assembly (Witte and Baltimore, 1978; Vogt et al., 1979): the virions do, however, bud in a similar manner.

The formation of virions has several functions. The most obvious one being the creation of a particle that is extracellular and capable of infecting other cells. The process of infection has two stages, adsorption to the cell surface and

penetration. Relatively is known about either but entry does require the assistance of plasma membrane receptors that interact specifically with glycoproteins in the viral envelope (reviewed by Varmus, 1983). Another consequence of virion formation is the juxtapositioning of the replicative machinery with its template (RNA). Reverse transcription takes place after penetration and produces the cDNA required for chromosomal insertion. The details of the reverse transcription reaction are complicated (Gilboa *et al.*, 1979). The primer is the tRNA that binds to the genomic RNA 3' to the 5' LTR, the synthesis initiates at this point and continues to the 5' end of the RNA (a strong stop in synthesis). This nascent DNA strand is transferred to the 3' end of the viral RNA, thus utilising the R region terminal redundancy of the RNA. For this to occur the RNA of the RNA-DNA duplex has to be removed by another *pol* encoded activity, RNaseH. Synthesis of a linear cDNA is then completed. The linear cDNA copy has a complete LTR at both ends, as well as two additional nucleotides that arise from the nucleotides that are between the internal junctions of both LTRs, and the primers for plus and minus strand synthesis (Figure 1.5; Varmus, 1985). Panganiban and Temin (1984) have shown that this linear molecule forms a double LTR circle that is the intermediate for integration into the host's genome (Figure 1.5). The LTR-LTR junction along with the four additional base pairs is contained within a 20 bp sequence called *att*: the junction is recognised by the integrase (a further cleavage product of *pol*) to produce a staggered nick that allows chromosomal integration of the provirus to occur with the concomitant loss of the extra nucleotides (Duyk *et al.*, 1985;

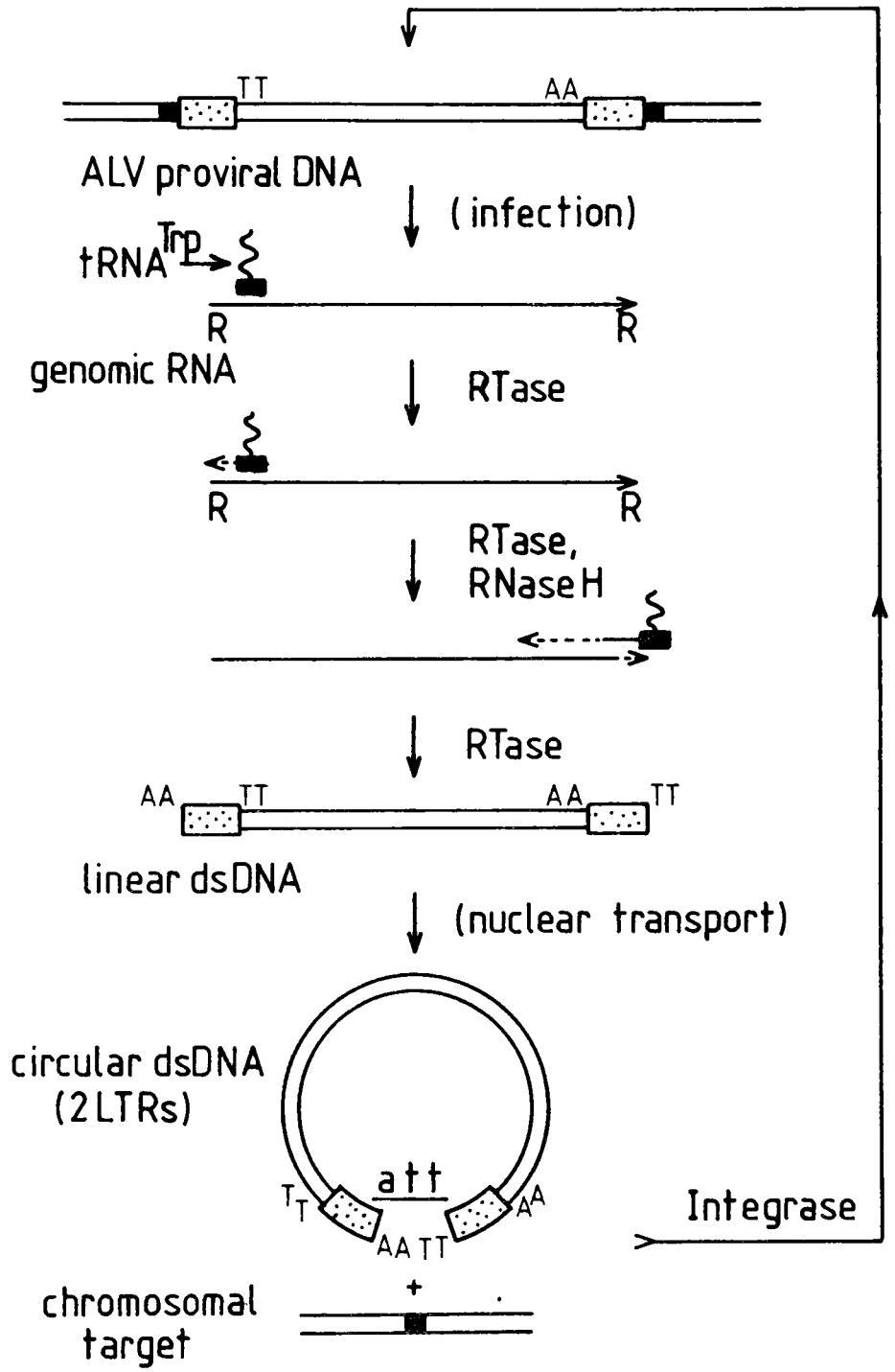


Figure 1.5. The life cycle of Avian Leukosis Virus. Solid boxes, genomic target sequence; stippled boxes, LTRs; double lines, DNA; single lines, RNA; R, terminal redundancy.

Grandgenett et al., 1986). Once inserted into the host's genome the provirus will be inherited in a Mendelian manner. Upon the necessary stimulus, it may start expressing RNA and protein so that the whole cycle can be repeated. It has recently been shown by Linial (1987) that the transmission of retroviral sequences from one cell to another does occur via an RNA intermediate, thus confirming this predicted life-cycle (Figure 1.5).

The life-cycle outlined above enables retroviruses to evolve rapidly. In other words their genomes may be regarded as plastic; integration into host genomes, the fidelity of reverse transcription and frequent recombination with other retroviruses all contribute to this plasticity. A good example of the recent emergence of a "new" retrovirus is HIV (reviewed by Gallo, 1987).

This discussion has been restricted to replication-competent retroviruses. These are often described as chronic retroviruses because pathogenesis is much slower than when acute retroviruses are the causative agent. There are several mechanisms by which chronic retroviruses are able to cause disease by over-expressing cellular oncogenes (c-onc) at a time when expression would normally be minimal or zero. Payne et al. (1982) have identified three configurations of integrated proviral ALV DNA that activated c-myc in bursal lymphomas (Figure 1.6). In all three cases the levels of c-myc are elevated several fold: Type I activation is known as promoter insertion since the highly expressing ALV 3'LTR overrides the c-myc promoter (note that the 5' LTR is deleted), Type II activation involves the ALV 3' LTR acting as both an enhancer and a transcription terminator, and

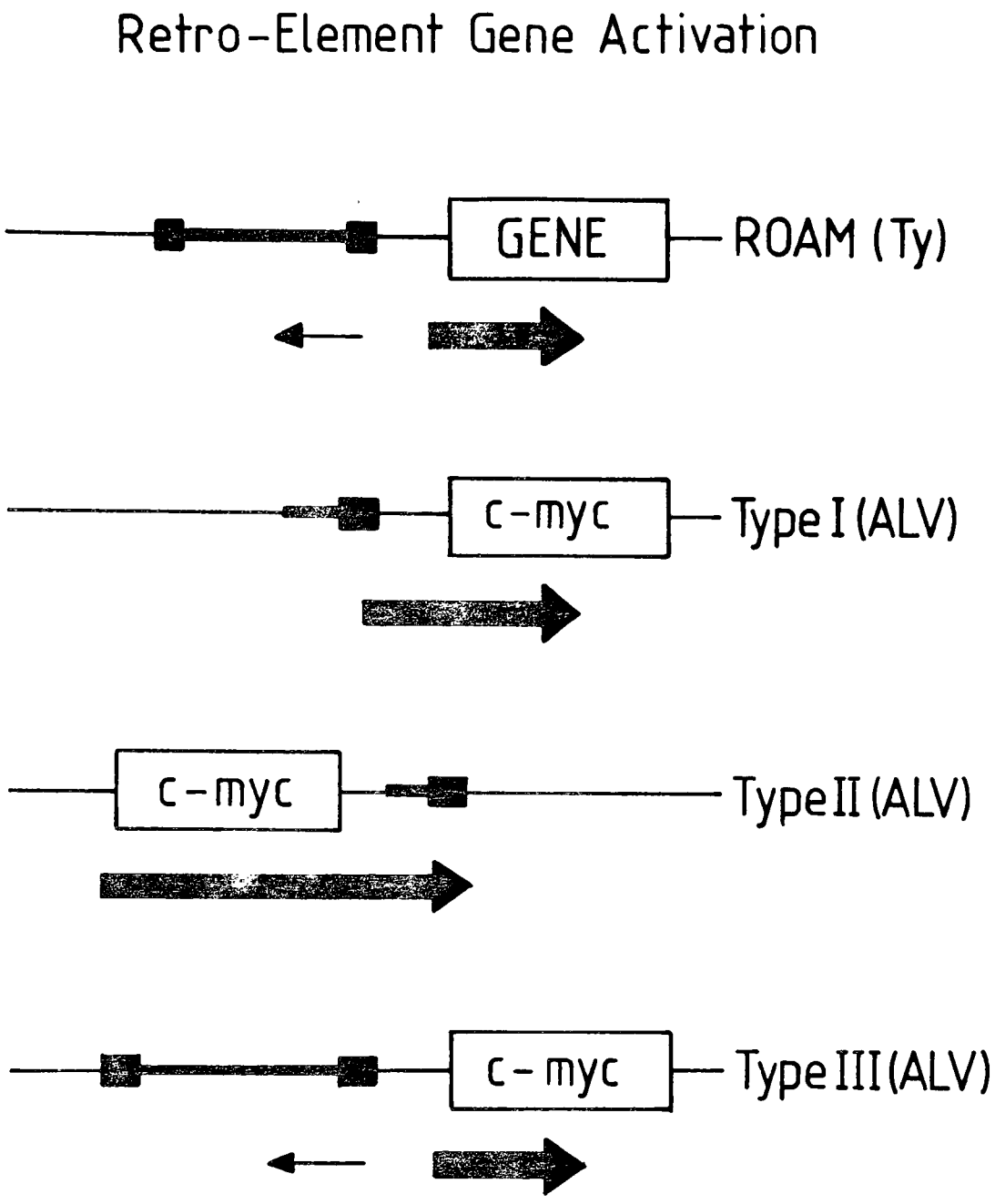


Figure 1.6. Retro-element gene activation. Open boxes, activated gene; solid boxes, LTRs; thick lines, retro-element unique regions; thick arrows, abundant transcripts of activated genes; thin lines, relatively rare transcripts of retro-elements. Adapted from Payne et al. (1982).

Type III activation utilises some type of enhancer element in ALV. In Type III activation, authentic c-myc RNA is still produced but at elevated levels whilst the directions of ALV and c-myc transcription are opposed.

Acute transforming retroviruses function rather differently and are usually replication-incompetent. They acquire and mutate c-onc genes to produce v-oncs, often with the loss of regions of their own retroviral genes; this frequently renders the virus incapable of supporting itself such that a helper, replication-competent retrovirus is necessary for its propagation and transmission. The mutated v-onc proteins are highly expressed under the control of the retroviral 5' LTR and may cause rapid malignant transformation as a consequence of their elevated levels or their novel, yet uncontrollable, activity within a cell. The products of v-onc genes are often protein kinases (usually tyrosine kinases) but may also be involved in the regulation of transcription or DNA replication (reviewed by Bishop, 1985).

1.3.3.2 Copia-Like Elements of Drosophila

About five percent of the D. melanogaster genome comprises 30-50 families of repetitive elements with very similar properties, these are known as the copia-like elements (reviewed by Rubin, 1983; Finnegan, 1985). The most studied families are copia, 412, 297, mdg1, mgd3, mdg4 (gypsy), B104 and 17.6. The details of copia itself will be discussed here.

There are approximately fifty genomic copies of copia. The

genome of one element from the white-apricot allele has been fully sequenced (Mount and Rubin, 1985). It has the common structural features of a retrotransposon (Figures 1.2 and 1.4). The element is 5,146 bp long and has LTRs of 276 bp. There is a single ORF of 4,227 nucleotides and this contains the consensus sequences for a retroviral acid protease, a reverse transcriptase, and an integrase as well as a bacterial nucleic acid binding protein (Mount and Rubin, 1985). The relative positions of the reverse transcriptase and integrase are reversed with respect to retroviruses (Figure 1.4). The various copia-like elements, have similar genomic organisations but their actual sequences are unrelated; 17.6 is rather larger (7,349 bp) than the other elements and has three ORFs, the most 3' one of which may correspond to retroviral env genes (Figure 1.4) (Saigo et al., 1984).

Copia-like elements are transcriptionally active to the extent that copia RNA may constitute three percent of the cytoplasmic polyadenylated RNA of Drosophila cells. A 5 kb full length transcript is produced that runs from the 5' LTR to the 3' LTR (Carlson and Brutlag, 1978); other smaller RNA species have also been identified, the most abundant being a 2 kb message. Both the 5 kb and 2 kb messages are polyadenylated at their 3' ends and capped at their 5' ends, though the start site of both RNAs is rather heterogeneous (Flavell et al., 1981). The signals for the transcription of these RNAs have been localised to the 5' LTR (Sinclair et al., 1986) but in contrast to retroviruses (Section 1.3.3.1), they are downstream of the RNA start site. Both the RNAs are translated inefficiently but a 51 kD protein

can be produced *in vitro* (Flavell *et al.*, 1980). The transcription of copia-like elements is developmentally regulated by genes at unlinked loci (Parkhurst and Corces, 1987). Consequently different elements are expressed at various stages of development: copia transcription peaks during the pupal stage, but also occurs in the embryonic and adult stages of the life cycle.

In 1983, Shiba and Saigo showed that Virus-Like Particles (VLPs) with a diameter of 50 nm were present in the nuclei of cultured *Drosophila* cells. These VLPs had associated copia RNA and reverse transcriptase. The major protein in the VLPs was 31 kD; a protein with a similar molecular weight, that was immunologically cross-reactive, was produced by the *in vitro* translation of the VLP RNA. This demonstrated that the VLPs contained copia encoded proteins. The discovery of copia-VLPs indicated that these repetitive elements were retrovirus-like with respect to the function of their proteins as well as their genome organisation.

However, it has not been shown that copia transposes via RNA but the strong stop products of reverse transcription of some copia-like elements have been identified in cultured *Drosophila* cells (Arkhipova *et al.*, 1986). Various species of linear and circular double stranded DNA molecules that contain copia sequences have also been found in *Drosophila* cells. Many of the circles are the likely products of homologous recombination between the LTRs of chromosomal elements (Flavell and Ish Horowicz, 1983); the linears are currently believed to be the

true products of reverse transcription (Flavell and Brierley, 1986). This will be discussed further in Section 8.1.

The genetic effects of copia and copia-like elements are well characterised (for the white gene refer to Levis et al., 1984; and for the forked gene refer to Kubli, 1986). Mutations at these loci that are due to transposon insertion can be reversed by mutations at unlinked loci; for example, mutations in Su(Hw) can restore the transcription of the forked gene, presumably by inhibiting the transcription of the gypsy element responsible for the initial mutation (Parkhurst and Corces, 1985). This demonstrates the close association of the host cell's status to retrotransposon expression. The mechanisms of these phenomena are currently being elucidated.

1.3.3.3 TED Elements of Trichopulsia ni

The TED retrotransposon is present in about fifty copies per genome of the moth Trichopulsia ni. They were discovered as a polymorphism in the genome of the baculovirus Autographica calcifornia (Miller and Miller, 1982) during an infection of T. ni cells in culture. TED has the structural features of a retrotransposon (Figure 1.2): the element is 7.3 kb long with LTRs of 270 bp, but it has not been fully sequenced as yet. A full length transcript (7 kb) is produced as well as a number of smaller RNA species (Friesen et al., 1986). The excision of TED by homologous recombination leaves a solo LTR that is transcriptionally active in both directions and promotes transcripts that extend into the flanking sequences (Friesen et

al., 1986).

1.3.3.4 Intracisternal A Particles of Rodents

The genomes of rodents such as mice, rats and Syrian hamsters contain approximately one thousand copies of a highly polymorphic repetitive DNA element that has homology to the RNA of the non-infective Intracisternal A Particles (IAPs) (Lueders and Kuff, 1977), that are found budding from the endoplasmic reticulum of various embryonic and tumour cell lines (Kuff et al., 1968). These elements are called IAP genomes and have a size range of 3-7.3 kb in Mus musculus, suggesting that a certain number are shortened and probably defective (rather like the P element heterogeneity in Drosophila). The sequencing of a full length element, IAP-H18, from Syrian hamster (Ono et al., 1985) revealed the typical LTR-gag-pol-env-LTR structure of retroviral proviruses (see earlier), consequently IAP genomes are classified as retrotransposons. IAP-H18 is 7,951 bp long with LTRs of 376 bp and has the consensus sequences of retroviral reverse transcriptases and integrases. The "gag" and "pol" genes are imperfect ORFs whilst the "env" region contains numerous stop codons. This genomic structure is reminiscent of the Drosophila retrotransposon 17.6 but rather different to the copia and TED elements described earlier (Figure 1.4). The LTRs are capable of driving efficient expression of heterologous genes in mouse cells (Lueders et al., 1984) and resemble the proviral LTRs in that the eukaryotic consensus sequences for transcription initiation and

polyadenylation are present (Ono et al., 1985).

The diameters of IAPs range from 70 to 100 nm (Kuff et al., 1968). They have an associated reverse transcriptase activity (Wilson et al., 1974) as well as IAP genome specific RNA (Lueders et al., 1977). The major protein of murine IAPs is 73 kD and is encoded by the IAP genome itself (Paterson et al., 1978). In addition, the IAP proteins are immunologically cross-reactive with the structural proteins of murine retroviruses (Kuff et al., 1980). The presence of stop codons in the "gag" and "pol" regions of IAP-H18 makes it difficult to see how it could produce authentic proteins; it is possible that this particular element is not expressed and is therefore unrepresentative. The budding process of IAPs results in their having a small amount of associated phospholipid, and may be reminiscent of retroviral particle formation at the plasma membrane. IAPs are completely non-infectious, and this might be due to the lack of an "env" protein. IAP genomes have been observed to transpose (Canaani et al., 1983; Kuff et al., 1983); this has been postulated , but not proven, to occur via an RNA intermediate.

1.4 The Yeast Retrotransposon Ty

The haploid genome of S. cerevisiae is 2×10^7 nucleotides in size. 85-95% of this genome is single copy sequence (Lauer et al., 1977). Most of the remainder comprises various middle-repetitive elements (though a small amount of highly-repeated sequence does exist, refer to Section 1.2). These repeated sequences include the rRNA, tRNA and histone genes as well as the

Ty, sigma (Delrey et al., 1982) and tau (Chisholm et al., 1984) elements.

1.4.1 General Features

Ty elements were first identified as the cause of polymorphisms at the SUP4 locus on chromosome X (Cameron et al., 1979), and since then several Ty elements have been cloned. Ty has the structural features of a retrotransposon (Figure 1.2). The further characterisation of various Ty elements has shown that they share many of the features, like consensus sequences, expression strategy and protein functions, that have already been described for copia, TED and IAP genomes as well as integrated retroviral proviruses (the retro-elements).

There are 30-35 copies of Ty in the haploid genomes of most laboratory strains of S. cerevisiae (approximately one percent of the genome) (Cameron et al., 1979; Kingsman et al., 1981), these are dispersed throughout the genome so that most of the chromosomes contain at least one Ty element. Restriction mapping and heteroduplex analysis has shown that Ty elements are polymorphic, but that two broad classes exist: there are approximtely 30 Class I elements and 5 Class II elements (Kingsman et al., 1981). Most Ty elements are 5.9 kb long with LTRs (or deltas) of 340 bp (Roeder and Fink, 1982). The 5.2 kb internal region is unique and is called epsilon. In addition to the full length copies of Ty there are approximately 100 solo-deltas present in the haploid genome (Cameron et al., 1979). These are probably the result of homologous recombination between

the delta sequences of a full length element. Both full length elements and solo-deltas are flanked by 5 bp duplications of the target site DNA (Farabaugh and Fink, 1980). There is a sequence at the 5' LTR-epsilon junction that is complementary to the 3' end of the yeast $tRNA^{met}$ (Eibel et al., 1980) and a conserved 5'-GGGTGGT-3' region at the 3' end of epsilon (Warmington et al., 1986), these may both have a role in reverse transcription.

Three Ty elements have been sequenced to date, Ty109 (Hauber et al., 1985), Ty912 (Clare and Farabaugh, 1985) and Ty1-17 (Warmington et al., 1986). Ty109 and Ty912 are Class I elements whilst Ty1-17 is a Class II element. The sequences have revealed a common genomic organisation for Ty elements (Figure 1.4). There are two ORFs that are designated TYA and TYB. For the class I element Ty109 (Hauber et al., 1985), TYA starts with an ATG (methionine) at residue 291 (within the 5' LTR) and extends to residue 1610; TYB starts with an ACA (threonine) at residue 1572 and extends to residue 5556. TYB therefore overlaps TYA by 38 bp and is in the +1 reading phase. Class II elements have the same oganisation except that the TYA:TYB overlap is 44 bp (Warmington et al., 1986).

The relationship of TYA and TYB is therefore reminiscent of the gag and pol genes of ALV (Figure 1.4). Searches for consensus sequences have revealed homologies to bacterial nucleic acid binding proteins in TYA (Pabo and Sauer, 1984) as well as to retroviral acid proteases (Toh et al., 1985), reverse transcriptases (Clare and Farabaugh, 1985; Hauber et al., 1985) and integrases (Mount and Rubin, 1985) in TYB. An obvious

difference between Ty and ALV is the absence of a region that could correspond to the env gene: in this respect, Ty resembles copia and differs from 17.6 and IAP genomes (Figure 1.4).

The major sequence differences between Class I and Class II elements have been mapped by heteroduplex analysis (Kingsman et al., 1981): they are clustered in TYA and the middle portion of TYB. The differences in TYA are due to extensive point mutations throughout the region. However, the predicted protein sequences of Ty1-15 and Ty1-17 are 48% homologous with a further 52% of the differences being conservative (Fulton et al., 1985). This maintenance of protein sequence (and, by implication, protein structure) suggests that proteins expressed from TYA have a conserved function.

1.4.2 Recombination between Ty and Delta Sequences

Since there are 30-35 copies of Ty and over 100 solo-deltas, the potential for homologous recombination between different parts of the genome is considerable (reviewed by Roeder and Fink, 1983). Transposition of Ty is a very different process, it is a non-homologous recombination event and will be discussed in Section 1.4.6. Recombination events between Ty elements or deltas may or may not rearrange the flanking markers; events which do not rearrange the flanking sequences may be reciprocal, whereby the sequences are exchanged, or non-reciprocal, whereby the sequences become the same (otherwise known as gene conversion). These various events presumably initiate by the alignment of homologous sequences, however the end results are

varied and include the excision, deletion, inversion or translocation of affected sequences. The frequency at which homologous recombination occurs is affected by unlinked loci. For example, the rate at which delta-delta recombination occurs so that Ty912 is excised from the HIS4 promoter (the his4-912 mutant) is increased 100 fold by spt1 and spt3 mutations and 1,000 fold by spt2 mutations. Moreover, this increased rate of excision in spt2 mutants is decreased 100 fold if the strain is also a rad52 mutant. The SPT (suppressor of Ty) genes have a profound effect on Ty transcription as well as affecting Ty-Ty recombination (see Section 1.4.3).

1.4.3 Transcription

Ty RNA constitutes 5-10% of the polyadenylated RNA of logarithmically growing yeast cells (Elder et al., 1980). The major species is a 5.7 kb transcript (the genomic RNA): this initiates at the U3-R junction of the 5'LTR (an XhoI site in Ty1-15) and terminates at the R-U5 junction of the 3'LTR (Figure 1.7) (Elder et al., 1983). The genomic RNA is therefore terminally redundant and analagous to transcript I of ALV (Figure 1.3). In cells that are heterozygous diploids (a/α), the level of Ty RNA decreases five fold (logarithmic cells) or twenty fold (stationary cells) (Elder et al., 1980). This phenomenon is known as mating-type control. Thus, Ty resembles other retrotransposons in that the host cell's status affects its level of transcription.

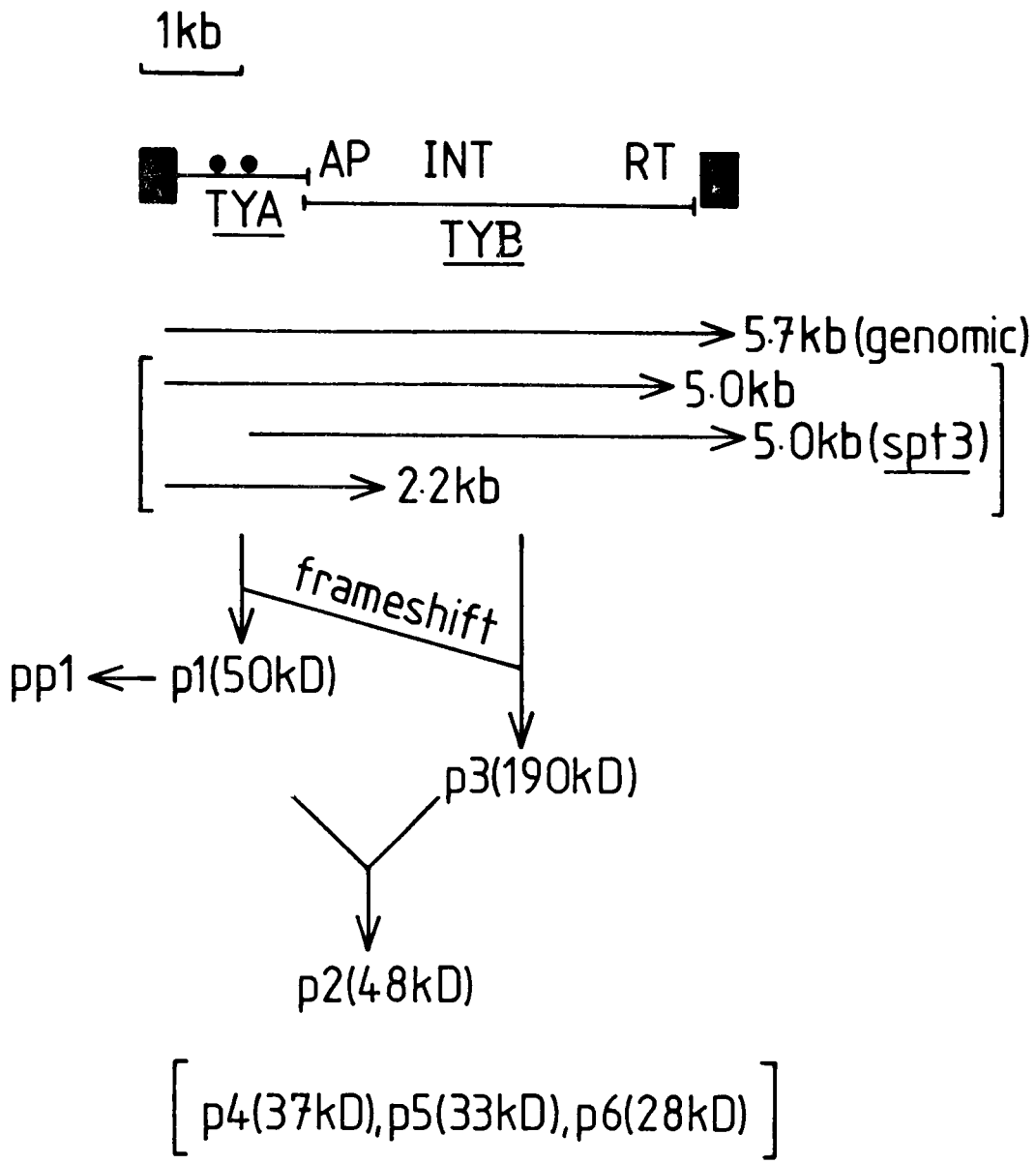


Figure 1.7. Genome organisation and expression strategy of a Ty element (1985). Solid boxes, LTRs; thin lines, ORFs; solid circles, homology to SV40 core enhancer (Weiher *et al.*, 1983) and sequences that respond to mating-type control signals (Siliciano and Tatchell, 1984); AP, INT and RT, homologies to the consensus sequences of retroviral acid proteases, integrases and reverse transcriptases; thin lines with arrowheads, RNA species; p, protein; pp, phosphoprotein.

The LTRs contain the consensus sequences of eukaryotic promoters and terminators (Bowen et al., 1984; Elder et al., 1983; Williamson et al, 1983), these include the conserved TATA (initiation) and TAGT (termination) components (Corden et al., 1980; Zaret and Sherman, 1982). A good candidate for the Ty promoter would therefore be the 5' LTR; this would be somewhat analogous to retroviral promoters (Section 1.3.3.1). However, it has been shown that there are positive elements for Ty transcription downstream from the RNA start site (Fulton et al., submitted); the deletion of these resulted in almost total abolition of 5.7 kb RNA transcription. If the LTR sequences upstream of the RNA start site (including the TATA box) are deleted, 5.7 kb transcription is again abolished (Fulton et al., submitted). Therefore, it appears that there are mutually dependent upstream and downstream elements involved in the regulation of transcription of the 5.7 kb species. This may explain the "promoter inactivity" of the 3' LTR. The sequences responsible for the mating-type regulation of Ty transcription reside within the 5' LTR, but downstream from the RNA start site (Malim et al., 1987a).

Other Ty transcripts have also been identified (Figure 1.7). A 5.0 kb species was identified by Elder et al. (1980) and said to be 5' coterminal with the 5.7 kb species; there is some doubt over the authenticity of this species since this data has never been confirmed. The major transcript in spt3 mutants is also a 5.0 kb RNA, however this has a start site about 1050 bp into the element (Winston et al., 1984b): the 5.7 kb species is virtually absent in these mutants. A small 2.2 kb RNA has also been

identified (Taguchi et al., 1984; Mellor et al., 1985c), the termini and transcription signals of which are unknown.

1.4.4 Genetic Effects

The integration of a Ty element at a new locus, by transposition, occurs preferentially into the promoter regions of genes (Eibel and Philippsen, 1984). The consequences of Ty insertion for gene expression can be very different (reviewed by Roeder and Fink, 1983). HIS4 transcription is prevented by the insertion of Ty912 (in the same transcriptional orientation) or Ty917 (in the opposite transcriptional orientation) into the HIS4 promoter, a phenomenon known as gene disruption. Variations in the Ty sequences in the his4-917 mutation resulted in phenotypes that were $his4^-$, weakly $his4^+$ (cold sensitive) and strongly $his4^+$ (overexpressing): both $his4^+$ derivatives displayed mating-type control. Mutations caused by the insertion of a complete Ty element, rather than by the alteration of a pre-existing one, have been identified that result in 20 fold activation, of CYC7 by overexpression (Errede et al., 1980b), and of ADR2 by constitutive overexpression (Williamson et al., 1981). These two activation events bring the genes under mating-type control so that there is no activation in a/α diploids. These mutants are called ROAM (Regulated Overproducing Alleles responding to Mating signals). In each case the direction of Ty transcription is away from the activated gene, thus Ty is acting as an "enhancer" rather than a promoter. The ROAM phenomenon does not require

transcription of the causative Ty element. The nature of the RNA species from the activated gene is unchanged in ROAM. ROAM is therefore directly analogous to the Type III activation of c-*myc* by ALV (Figure 1.6; Payne *et al.*, 1982).

The sequences of TyH2 responsible for the ROAM activation of *CYC7*-H2 do not reside in the LTRs and have been localised to between residues 573 and 1300 (Errede *et al.*, 1984; Errede *et al.*, 1985). The activators are likely to be in two regions, one from 730 to 770 and the other from 910 to 950. Both regions contain homology to the SV40 core enhancer (Weiher *et al.*, 1983) and a sequence known to respond to mating-type control signals (Siliciano and Tatchell, 1984). Results that conflict with Errede's observations have been obtained using *in vitro* ROAM mutations whereby the Upstream Activating Sequence (UAS) of the yeast *PGK* gene has been replaced with various Ty fragments (Rathjen *et al.*, submitted). Sequences sufficient for both activation and mating-type control have been found between residues 389 and 480 (i.e., considerably 5' to those of Errede *et al.*, 1985); these sequences are only active when in the 'ROAM' orientation.

Selection for the suppressors of Ty and delta insertion mutations at *HIS4* and *LYS2* has identified 12 trans-acting *SPT* genes (Winston *et al.*, 1984a; referenced in Clark-Adams and Winston, 1987). The *SPT3* gene is essential for normal transcription of the 5.7 kb RNA (delta mediated transcription) (see Section 1.4.3). An *spt3* null mutant results in the virtual loss of the 5.7 kb RNA and an increase in the levels of a 5.0 kb RNA (Winston *et al.*, 1984b). Overexpression of the *SPT3* protein

does not affect the levels of Ty RNA, therefore this protein is not the only limiting factor required for transcription. It seems unlikely that the gene product of the SPT3 gene interacts directly with the Ty promoter sequences since no DNA binding activity has been detected for the SPT3 protein (Winston and Minehart, 1986). The SPT6 gene product has been described as being essential for delta-mediated, but not whole element, transcription (Clark-Adams and Winston, 1987). It is hard to reconcile this conclusion with the data since transcription from the solo delta in his4-912δ is increased in an spt6 mutant. Therefore, the modes of action of the SPT proteins in Ty transcription are largely unelucidated.

As well as the mating-type control, a variety of unlinked genes, apart from SPT, suppress the ROAM phenomenon in haploids. Such mutants include ste7, ste11, ste12 (Errede et al., 1980a; Errede et al., 1980b; Errede et al., 1984), roc1, roc2 (Dubois et al., 1982) and tye (Ciriacy and Williamson, 1981). These effects resemble those seen with other retrotransposons and retroviral proviruses (Section 1.3.3.1), for example, the reversion of the gypsy induced mutation of forked in Drosophila by a mutation in Su(Hw) (Parkhurst and Corces, 1985).

1.4.5 Ty Proteins and the Strategy of Overexpression

Ty proteins have been identified by the overexpression of Ty ORFs in yeast expression vectors. The most successful vector that has been used is pMA91 (Mellor et al., 1983; Kingsman and

Kingsman, 1983). It contains a BglII expression site between the powerful yeast PGK promoter and the PGK terminator (Figure 1.8). pMA91 is present at approximately 100 copies per cell as a result of the LEU2/2μ selection and replication module, and has been used to express interferon and calf chymosin in yeast. The TYA region of Ty1-15 was inserted, as a manipulated BamHI-BglII fragment (the 11 fragment), into the expression site of pMA91 to yield pMA91-11 (Figure 1.8) (Dobson et al., 1984). Extracts from yeast cells containing pMA91-11 contained a protein of 50 kD that was not present in control extracts, this was designated p1(1-15) (Dobson et al., 1984). The whole Ty1-15 transcriptional unit (TYA and TYB regions) was manipulated as a BamHI fragment (the 10 fragment) and inserted into pMA91 to yield pMA91-10 (Figure 1.8) (Dobson et al., 1984). Two further Ty proteins were identified in a pulse-chase experiment using pMA91-10 transformed cells. In the pulse, the 50 kD protein, p1(1-15), and a 190 kD protein, p3(1-15), were present; in the chase a 48 kD protein, p2(1-15), appeared with the concomitant decrease of p1 and disappearance of p3 (Mellor et al., 1985b). It is interesting that the appearance of p2 was dependent upon the expression of TYB, and that TYB contains the consensus sequence for retroviral proteases (Figure 1.4) (Toh et al., 1985); this suggests that a Ty encoded protease, rather like retroviral proteases, may mediate the proteolytic processing of the primary translation products of Ty. It has been shown that p1 is a phosphoprotein and that both p1 and p2 bind DNA in a crude binding assay (Mellor et al., 1985b). Few yeast proteins have this property, so that the smaller Ty proteins p4 (37 kD), p5 (33 kD) and p6 (28 kD) have been

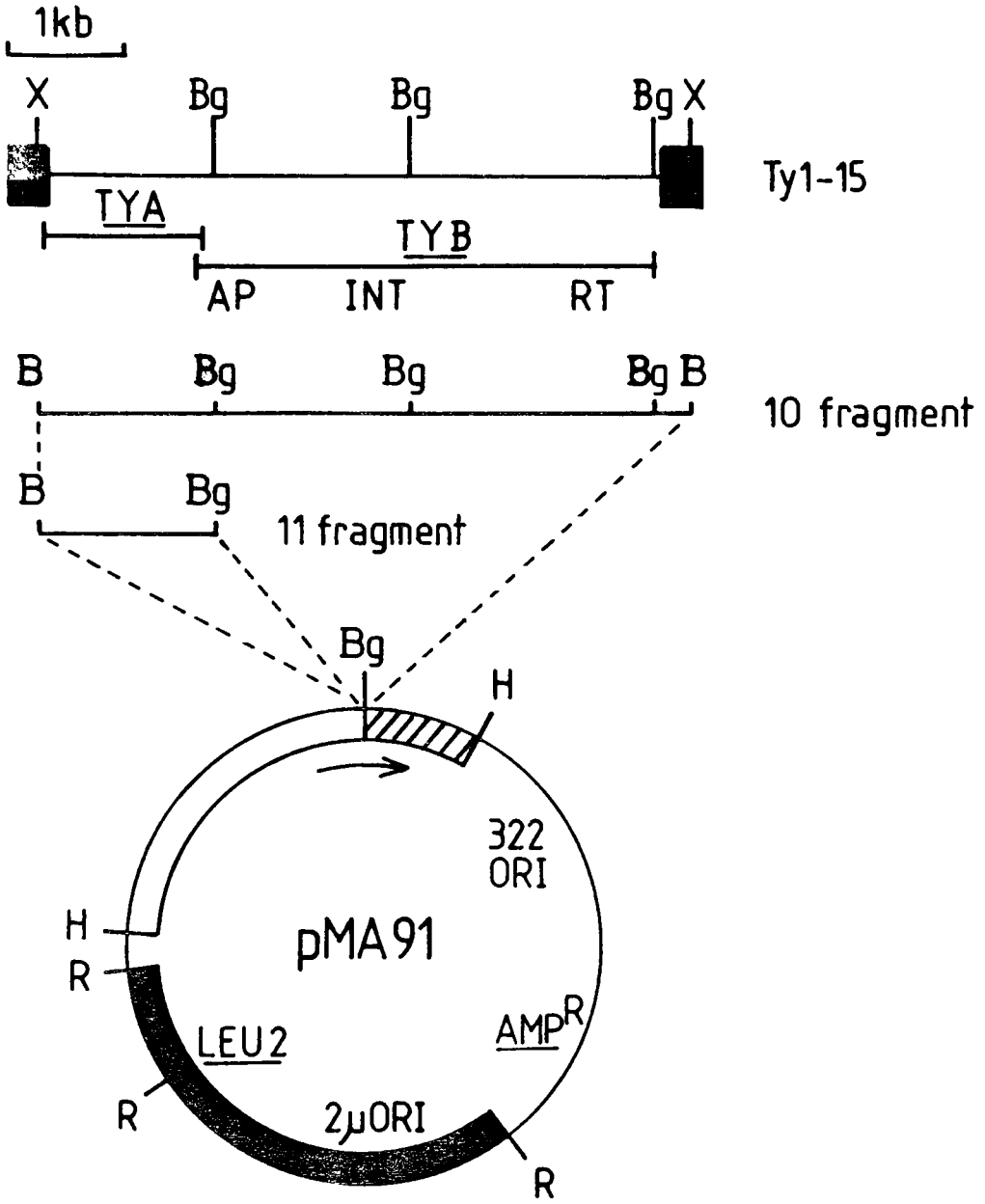


Figure 1.8. The insertion of Ty1-15 ORFs into pMA91. The genome organisation of Ty1-15 is described in the legend to Figure 1.7. Thin lines, pBR322 derived sequences; solid box, LEU2/2μ origin selection and replication module from pJDB219 (Beggs, 1978); open box, 5' promoter region from the yeast PGK gene; hatched box, 3' terminator region from the yeast PGK gene. The arrow indicates the direction of transcription promoted by the PGK 5' region. The BamHI fragments, 10 and 11, were inserted into the BglII expression site of pMA91 to yield the plasmids pMA91-10 and pMA91-11, respectively. X, XhoI; Bg, BglII; B, BamHI; R, EcoRI; H, HindIII.

identified by virtue of their DNA binding activity (Malim et al., 1987a). The use of pMA91 was probably influential towards the identification of Ty proteins because of the utilisation of the PGK RNA start site as well as the increased levels.Thus, 40 nucleotides of PGK leader RNA were present upstream of the authentic Ty RNA start site: this sequence may contain a strongly positive translation signal since it is known that the PGK protein is normally expressed efficiently (Ogden et al., 1986). It is likely that, rather like copia RNA, Ty RNA is poorly translated in vivo.

TYA is expressed by simple translation so that the initiation codon for p1 is the ATG (methionine) at residue 291. TYB and TYA overlap by 38 bp, TYB being in the +1 reading phase with respect to TYA (Section 1.4.1). TYB is expressed as a TYA:TYB fusion protein (Mellor et al., 1985a; Clare and Farabaugh, 1985; Wilson et al., 1986); the 190 kD p3 protein is likely to be the in vivo TYA:TYB fusion protein. It has been shown that RNA splicing is not responsible for the expression of TYB (Wilson et al., 1986) and that a mechanism involving ribosomal frameshifting is a more likely explanation, with a region of 31 nucleotides, within the overlap region, being sufficient (Wilson et al., 1986).

The expression of TYA and TYB is analogous to the expression of the gag and pol genes of avian retroviral proviruses. TYA and gag are expressed by simple translation whilst TYB and pol are expressed as fusion proteins. However, the precise mechanisms of frameshifting are likely to be different. TYB is in the +1

reading phase with respect to TYA but pol is in the -1 reading phase with respect to gag.

1.4.6 Transposition

Boeke et al. (1985) have studied Ty transposition using a galactose inducible, GAL1 expression system. TyH3 expression was therefore controlled by the GAL1 promoter. Galactose induction resulted in the increased transposition of both TyH3 and the 35 endogenous Ty elements. TyH3 was marked with an intron-containing fragment. Transposition resulted in the removal of the intron, this demonstrated that Ty elements transpose via an RNA intermediate, a process known as retrotransposition. This is analagous to the spread of retroviral sequences via RNA (Linial, 1987). The transposition of Ty elements therefore requires reverse transcriptase; the mutagenic effect of transposition upon Ty elements is probably a result of the infidelity of this enzyme (Boeke et al., 1985). Any gene or gene product that affects Ty transcription is likely to affect transposition by changing the levels of the RNA intermediates. Transposition is reduced in spt3 mutants (Boeke et al., 1986) and a/α diploids (as a result of mating-type control).

Why yeast cells should have developed the ability to regulate the transcription, and therefore the transposition, of Ty elements is not clear. It may well be that such events are undesirable in certain situations and would only serve to reduce a yeast population's chances of survival during those times.

1.5 Summary and Thesis Objectives

It has been described how a range of eukaryotic transposons with Long Terminal Repeats (LTRs), known as retrotransposons (Boeke et al., 1985), possess a number of similarities to integrated retroviral proviruses. Transposons in this class include the Ty elements of S. cerevisiae, the copia-like elements of Drosophila, the IAP genomes of some rodents and the TED elements of T. ni.

All these elements have ORFs, some of which may be overlapping, as in the case of Ty (Figure 1.3). The retrotransposons that have been sequenced have the consensus sequences that proviruses contain; for example, homologies to bacterial DNA binding proteins, an acid protease, a reverse transcriptase and an integrase have been identified. The elements are transcriptionally active and produce a full length (genomic) RNA species that is terminally redundant. This has been shown to be the nucleic acid intermediate of retroviral infection and Ty transposition. The regulation of transcription appears to be different for the different elements, for example, retroviruses such as RSV can drive efficient transcription from the LTR alone; Ty appears to require signals that are located in the LTR and epsilon regions, whilst copia only requires signals that are in the LTR but downstream from the RNA start site.

Expression strategies are also shared. The genomic RNA is translated, often inefficiently, by utilising the first available ATG: however, the expression of ORFs other than the first ORF

requires alternative strategies. Ty and avian retroviruses (for example, RSV) express TYB and pol, respectively, by a ribosomal frameshifting event that brings the out of phase first and second ORFs into phase so that a fusion protein is produced. The functions of the proteins also show some uniformity: virus-like particles (VLPs) that are composed of element-encoded protein, element encoded RNA and possess a reverse transcriptase activity have been identified for copia elements in Drosophila nuclei and IAP elements in certain rodent cell lines. There is no evidence that these VLPs are infectious in the same way that retrovirus particles are. This may be due to the lack of an env equivalent (Figure 1.3), copia has no coding capacity for an "env" gene, and although IAP genomes do possess the space for an "env" gene, this region is full of stop codons (at least in IAP-H18).

Ty elements have been shown to produce proteins, and there is evidence of post-translational processing. The levels of the primary translation product of TYA (p1) have been shown to diminish if TYB is expressed, this results in the accumulation of a smaller protein, p2. A large precursor protein, p3, has been identified and this may be analogous to the gag-pol fusion protein of retroviruses. Rather like retroviral gag proteins, phosphorylation of p1 does occur. Therefore Ty elements have many of the characteristics of retroviral proviruses and retrotransposons. Further studies might reveal how Ty elements function, and indeed, if Ty does turn out to be a model retro-element, any information that is obtained might increase the understanding of the expression, properties, function or modes of action of the other retro-elements.

At the beginning of this thesis there were three main aims:

1. Do the previously identified Ty proteins have any function?

Two obvious properties to look for were the ability to form particles (VLPs) and a reverse transcriptase activity. There was good reason to think that Ty proteins would have these functions because the proteins encoded by both copia and IAP elements form VLPs that have associated reverse transcriptase activity.

2. Are there any DNA intermediates of Ty transposition?

If Ty transposition is truely analagous to the spread of retroviral sequences, then circular double stranded DNA molecules with 2 LTRs should exist. Similar species have been identified in retrovirally infected cells and Drosophila cells (copia circles).

3. Does the expression of a Class II Ty element resemble that of a Class I element?

The search for a TYA:TYB fusion protein analogous to that produced by Class I elements would demonstrate the conservation of expression strategy and possibly function.

CHAPTER 2. MATERIALS AND METHODS

2.1 *E. coli* Strains

AKEC28 = C600, *thrC*, *leuB6*, *thyA*, *trpC1117*, *hsdRk*, *hsdMk*, *Str*^{R}.

JM103 = *dlacpro*, *thi*, *strA*, *endA*, *sbcB15*, *hsdR4*, *supE*, *F'traD36*, *proAB*, *lacI*^{q}.

HB101 = *F*$^{-}$, *hsdS20*($r^{-}_{B}M^{-}_{B}$), *supE44*, *ara14*, *galK2*, *lacY1*, *proA2*, *rpsL20*(str^{R}), *xyl5*, *leu*, *mtl1*, λ^{-}, *recA13*.

2.2 Yeast Strains

MD40-4c = *alpha*, *ura2*, *leu2-3*, *leu2-112*, *his3-11*, *his3-15*, *trp1*.

2.3 *E. coli* Media

Complete medium (L-Broth):

Tryptone, Difco	10 g
yeast extract, Difco	5 g
NaCl	5 g
thymidine	50 mg
H_2O	1 litre
agar (for solid medium)	20 g

Sterilise by autoclaving, then add, when required:

antibiotics	25 mg

Minimal Medium (EMM):

K_2HPO_4	7.0 g
KH_2PO_4	3.0 g
Na citrate	0.5 g
$MgSO_4$	0.1 g
$(NH_4)_2SO_4$	1.0 g
glucose	2.0 g
agar (for solid medium)	20.0 g
H_2O	954 ml

Sterilise by autoclaving.

The specified volumes of the following sterile solutions are added as required:

Ø.4% threonine	1Ø ml
Ø.5% thymine	1Ø ml
thymidine (2.5 mg/ml)	2Ø ml
Ø.2% thiamine	1 ml
Ø.2% tryptophan	1Ø ml
Ø.3% leucine	1Ø ml
antibiotics	25 mg

2.4 Yeast Media (Fink, 1970)

Complete medium (YPD)

yeast extract	1Ø g
peptone	1Ø g
glucose	2Ø g
H_2O	1 litre
agar (for solid medium)	2Ø g

Sterilise by autoclaving.

Synthetic complete medium (SC)

yeast nitrogen base (without amino acids)	6.7 g
glucose	1Ø.Ø g
H_2O	1 litre
agar (for solid medium)	25.Ø g

Sterilise by autoclaving, then add:

Mixed amino acid solution	2Ø ml
Ø.3% leucine (if required)	1Ø ml
Ø.2% tryptophan (if required)	1Ø ml

Mixed amino acid solution:

adenine	Ø.2 g	isoleucine	3.Ø g
histidine	Ø.2 g	phenylalanine	Ø.5 g
uracil	Ø.2 g	valine	1.5 g
tyrosine	Ø.3 g	threonine	2.Ø g
lysine	Ø.3 g	serine	3.5 g
arginine	Ø.2 g	glutamic acid	1.Ø g
methionine	Ø.2 g	aspartic acid	1.Ø g

2.5 Plasmid DNA isolation from E. coli

2.5.1 Rapid Procedure ("miniprep") (Holmes and Quigley, 1981)

1. Grow 1 ml of culture in L-Broth, shaking at 37°C.
2. Harvest the cells in an eppendorf microfuge at room temperature (RT).
3. Resuspend the pellet in 350 µl of STET buffer.
4. Add 20 µl of 15 mg/ml lysozyme (in H_2O).
5. Place in boiling water for 45-60 seconds.
6. Centrifuge at 4°C for 10 minutes and remove the pellet with a toothpick.
7. Extract once with 100 µl of phenol/chloroform/isoamyl alcohol (25/24/1). The phenol was previously equilibriated against STE.
8. Add an equal volume of isopropanol, mix, and leave at -20°C for 10 minutes.
9. Centrifuge at 4°C for 5 minutes.
10. Remove the supernatant and resuspend the pellet in 10 µl of H_2O. Use about 2 µl of the resuspended DNA for restriction analysis on agarose minigels or about 5 µl for polyacrylamide gel analysis. The amount of DNA used should be empirically determined.

STET buffer:	sucrose	8 %
	Triton X-100	5 %
	EDTA (pH8.0)	50 mM
	Tris-HCl (pH8.0)	50 mM
STE:	NaCl	10 mM
	EDTA (pH8.0)	1 mM
	Tris-HCl (pH8.0)	10 mM

2.5.2 Large Scale Plasmid Preparation: Triton Method (Clewell and Helinski, 1969)

All plasmids used for DNA manipulation were prepared by this method.

1. Grow 100 ml of culture in a 250 ml flask, with vigorous aeration at 37°C, for about 24 hours.
2. Harvest the cells by centrifugation at 5 krpm for 10 minutes at 0°C, in a Beckman J2-21 centrifuge (rotor JA-10).
3. Wash the cells with H_2O and resuspend in 1.2 ml of cold 25% sucrose, 0.05M Tris-HCl (pH8.0) in a polypropylene tube.
4. Add 0.4 ml of freshly made lysozyme (10 mg/ml in H_2O). Swirl on ice for 10 minutes.
5. Add 0.4 ml of cold 0.5M EDTA (pH8.5). Swirl on ice for 5 minutes. EDTA removes Ca^{2+} from the cell wall to allow greater accessibility for the lysozyme.
6. Add 2.0 ml of cold triton solution (0.1% Triton X-100, 62.5 mM EDTA (pH8.5), 20 mM Tris-HCl (pH7.5)). Swirl on ice for at least 15 minutes. This should lyse the lysozyme treated cells. Vigorous shaking should be avoided as it releases large chromosomal DNA from the cell debris.
7. Centrifuge at 15 krpm for 60 minutes at 0°C (JA-17 rotor). The pellet should be gelatinous and the supernatant luminescent. Transfer the supernatant to a glass extraction tube.
8. Extract the supernatant with 1 ml of phenol/chloroform/IAA. Separate the phases by centrifuging at 9 krpm for 10 minutes at 0°C. Use a 15 ml corex tube (Dupont Instruments) for the separation.

9. Transfer the supernatant to a 30 ml corex tube and add 0.4 ml (1/10 volume) 3M sodium acetate and 10 ml (2.5 volumes) ethanol. Mix well and leave at -70°C for 30 minutes.

10. Centrifuge at 9 krpm for 60 minutes at -10°C. All the nucleic acids (DNA and RNA) should be in a rather mobile yellow pellet. Pour off the supernatant.

11. Resuspend the pellet in 7.7 ml STE.

12. Add 7.2 g of CsCl (Beckman, enzyme grade) and 0.3 ml of ethidium bromide (10 mg/ml in H_2O). The initial density of this solution is slightly higher than 1.50 g/cm^3. Transfer this solution to a Beckman Quick Seal polyallomer tube and top up with liquid parafin. Seal the tubes by melting.

13. Centrifuge at 45 krpm for 20 hours at 20°C. (Beckman rotor, Ti50). The CsCl gradient is established within a couple of hours but the distance the DNA molecules must travel, to reach their equilibrium positions, is large.

14. Visualise the DNA bands with a long wave (365 nm) ultraviolet lamp, model B-100 A/X (Ultra Violet Products, Inc., California). Release the seal at the top of the tube and remove the lower plasmid band (covalently closed circles) with an 18-gauge needle attached to a 1 ml syringe. The upper DNA band consists of open circles of plasmid DNA and sheared chromosomal DNA. Transfer the harvested band to a 15 ml corex tube.

15. Extract the ethidium bromide with three isopropanol extractions. The isopropanol should be equilibriated with NaCl saturated STE.

16. Add 2.0 ml of sterile H_2O and 6.0 ml ethanol (2.5 volumes).

Mix and leave at -70°C for 30 minutes.

17. Centrifuge at 9 krpm for 60 minutes at -10°C (JS13 swinging bucket rotor).

18. Pour off the supernatant and invert the tubes so as to dry the white pellets.

19. Resuspend the pellet overnight (RT) in 200 µl of STE.

20. Transfer the DNA solution to an eppendorf tube and add 0.5 ml ethanol. Mix and leave at -70°C for 30 minutes.

21. Centrifuge for 15 minutes at 4°C. Remove the supernatant and wash the creamy-white pellet with 70% ethanol (30% STE).

22. Resuspend the pellet in 50 µl of STE. The final concentration of plasmid DNA should be 1-2 µg/ul.

2.5.3 The Direct Transfer of Plasmid DNA onto Nitrocellulose Filters (Grunstein and Hogness, 1975)

This method was used to prepare filters for the rapid identification of specific insert sequences in plasmid DNA (colony hybridisation).

1. Place a 0.45 µm nitrocellulose circle (BA85/20, Schleicher and Schuell) onto a selective L-broth agar plate. On top of this, patch out the E. coli colonies to be screened (approximately 100 patches to a filter). Incubate the plates for 6 hours at 37°C.

2. During the 6 hour incubation, prepare 1 cm high, circular stacks of Whatman 3MM in four separate 15 cm diameter petri dishes.

3. Wet the 3MM in the dishes with the following:

   A. 0.2M NaOH

B. Ø.5M Tris-HCl (pH7.4)

C. Ø.5M Tris-HCl (pH7.4)

D. Ø.5M Tris-HCl (pH7.4), 1.5M NaCl.

4. After the 6 hour incubation, remove the n/c and place it (colony side uppermost) onto A. This should lyse the cells. Leave for 8 minutes.

5. Transfer the n/c to B and C, each for 2 minutes.

6. Transfer the n/c to D and leave for 4 minutes.

7. Dry the n/c on 3MM at 37°C for 1Ø minutes and then bake at 8Ø°C for 2 hours, under vacuum. These filters can be stored indefinitely and then probed for specific sequences at any time (see 2.18.6 for hybridisation protocol).

2.6 Plasmid DNA Preparation from Yeast

2.6.1 Rapid Isolation of Plasmid DNA from Yeast (Nasmyth and Reed, 1980)

This method was used to "rescue" plasmids from yeast.

1. Obtain a cell pellet (for example, the cells from a 5Ø ml selective culture grown to a density of 4 x $1Ø^6$ cells/ml) and wash once with H_2O.

2. Resuspend the cells in 1ØØ µl of 1M sorbitol, 1ØØ mM sodium citrate (pH7.Ø), 6Ø mM EDTA (pH8.Ø).

3. Add 1Ø µl of zymolyase 2Ø,ØØØ (Seikagaku Kogyo Co. Ltd., Japan) (1Ø mg/ml in H_2O) and 1 µl B-mercaptoethanol (BME) (14.3M). Incubate at 37°C for 3Ø minutes to spheroplast the

cells (removal of the cell wall).

4. Add 400 µl of 0.5% SDS, 100 mM Tris-HCl (pH9.7), 50 mM EDTA (pH8.5). Mix well and incubate at 70°C for 20 minutes.

5. Cool the mixture to RT and add 100 µl of 5M potassium acetate. Mix and leave on ice for 45 minutes.

6. Centrifuge for 5 minutes at 4°C in a microfuge.

7. Add 2 volumes of ethanol to the supernatant and place at -70°C for 20 minutes.

8. Centrifuge for 15 minutes at 4°C. Dissolve the pellet in 500 µl of STE and add 25 µl of 5M NaCl. The pellet does not fully dissolve but heating to 70°C for 5 minutes will help.

9. Phenol/chloroform/IAA extract twice.

10. Add 2 volumes of ethanol and place at -70°C for 20 minutes.

11. Centrifuge for 10 minutes at 4°C and resuspend the pellet in 50 µl of STE. Use 5 µl of the resuspended DNA for an *E. coli* transformation.

2.6.2 Large Scale Preparation of Plasmid DNA from Yeast (Cryer *et al.*, 1975)

This method is more usually used to isolate relatively pure chromosomal DNA, but may also be used for large scale plasmid isolation.

1. Grow 500 ml of cells to the desired density in the desired media.

2. Harvest the cells by centrifugation at 5 krpm for 10 minutes at RT (JA10 rotor).

3. Wash the cells with H_2O and transfer to polypropylene tubes,

pellet the cells.

4. Resuspend the cells in 2.0 ml of 45 mM EDTA (pH8.0), 300 mM BME and leave for 30 minutes at RT.

5. Harvest the cells by centrifugation at 9 krpm for 10 minutes at RT (JA-17 rotor) and resuspend in 2.0 ml of 1.2M sorbitol, 25 mM EDTA (pH8.0), M/15 sodium phosphate buffer (pH7.5).

6. Add 10 µl of zymolyase 20,000 (10 mg/ml in H_2O) and incubate at 37°C for one hour. This should spheroplast more than 90% of the cells.

7. Wash the spheroplasts once with the sorbitol/EDTA solution. This should be done gently as the spheroplasts are fragile (centrifugation is at 3 krpm at RT). The spheroplast pellet may be stored overnight at 4°C, if desired.

8. Resuspend the spheroplasts in 2.0 ml of 150 mM NaCl, 100 mM EDTA (pH8.0).

9. Add 10 µl of proteinase K (Sigma) (10 mg/ml in H_2O).

10. Add 50-100 µl of 25% SDS. Mix by gentle swirling and incubate, with occasional swirling, at 37°C for 3 hours. This should lyse the spheroplasts and release the DNA (less care is required for the preparation of plasmid DNA).

11. Heat to 60°C for 30 minutes, cool to RT, and transfer to a 15 ml corex tube.

12. Deproteinise by adding an equal volume of chloroform/IAA (24/1) and shaking gently until an emulsion is obtained. Separate the phases by centrifugation at 9 krpm for 10 minutes at RT.

13. Remove the aqueous phase and add 2 volumes of ethanol.

Place at -20°C for 1-2 hours.

14. Centrifuge at 9 krpm for 10 minutes at -10°C. A large, stringy pellet should be obtained. Resuspend the pellet overnight in 7.7 ml of STE. For plasmid DNA preparations, additional plasmid may need to be added as a "marker" in the CsCl gradient.

15. Add 7.2 g of CsCl and 0.3 ml of ethidium bromide and load into a Beckman polyallomer tube (see 2.5.2).

16. Centrifuge at 38 krpm for 36-48 hours at 15.4°C.

17. Visualise the DNA bands as before and remove the lower plasmid band.

18. Treat the plasmid DNA in the same way as plasmid DNA from *E. coli* (see Section 2.5.2, steps 14-22).

2.7 Gel Electrophoresis of DNA

2.7.1 Agarose Gel Electrophoresis

Horizontal agarose gels were routinely used for DNA analysis. Agarose concentrations from 0.8% to 1.5% were used depending on the size of the DNA fragments being studied, this should be empirically determined.

10X TBE:	Tris-base	108.0 g
	EDTA	9.3 g
	boric acid	55.0 g
	H_2O	1.0 litre

10X Tris-acetate:	Tris-base	193.6 g
	sodium acetate (anhydrous)	65.6 g
	EDTA	14.88g
	H_2O	4.0 litres
	pH to 8.2 with glacial acetic acid.	
10X gel loading buffer:	sucrose	10.0 g
	EDTA (pH8.0)	50.0 mM
	bromophenol blue	20.0 mg
	SDS	0.1 %
	H_2O (final volume)	20 ml

1X TBE was used for small scale gel ("minigel") electrophoresis, whilst 1X Tris-acetate was used for large scale electrophoresis (usually a large number of samples, requiring accurate sizing). Typical running times were 60 minutes at 110mA (130V) for a minigel (in TBE) and 3 hours at 250mA (300V) for a "large" gel (in 1X Tris-acetate). These conditions correspond to submerged gel electrophoresis.

To visualise the DNA, gels were submerged for 10 minutes in an ethidium bromide bath (1 μg/ml in H_2O), rinsed in H_2O and placed on a Chromo-blue Transilluminator TM-20 of Ultra-Violet Products, Inc., California (302 nm wave length). Photographs were taken with a Polaroid MD-4 land camera with an f2.8 lens aperture. Exposure was for 10-15 seconds using a Kodak 22-A red-orange filter. Polaroid, positive/negative film (Type 665, ASA 75) was used.

2.7.2. Polyacrylamide Gel Electrophoresis

Polyacrylamide gels were used for the analysis of DNA that

had been restricted many times to yield fragments with a size range of 40-600 bp.

30% acrylamide solution:	acrylamide BDH-Electron	29 g
	bis-acrylamide	1 g
	H_2O (final volume)	100 ml

acrylamide solutions were stored at 4°C in the dark.

7.5% acrylamide gel:	10X TBE	4.5 ml
	30% acrylamide	11.25 ml
	H_2O	29.0 ml
	10% ammonium persulphate	0.29 ml
	degas and add 30 µl TEMED	
2X gel leading buffer:	glycerol	1.0 ml
	10X TBE	0.5 ml
	bromophenol blue	2.5 mg
	xylene-cyanol-FF	2.5 mg
	H_2O (final volume)	5.0 ml

Use 1X TBE as the running buffer. A typical running time was 2 hours at 40 mA (200V). To visualise the DNA, gels were stained with ethidium bromide for 20 minutes and photographed in the same way as agarose gels.

2.8 Recovery of DNA Fragments from Gels

2.8.1 DNA Fragments from Agarose Gels

Fragments isolated by the following two methods can be used for DNA manipulation (Section 2.9) or DNA labelling (by nick translation, Section 2.18.5).

2.8.1.1 Using Glycerol

1. Separate the DNA fragments in an agarose gel containing ethidium bromide (1 µg/ml).
2. Visualise the DNA bands with a hand held long wave (365 nm) ultraviolet lamp (B-100 A/X Black Ray). Turn the current off. Remove the running buffer from the top of the gel (if submerged).
3. Using a scalpel, cut a thin well in the gel just to the positive side of the band.
4. Fill the well with a solution of 50% glycerol, 0.5X Tris-acetate (final concentrations).
5. Turn the current on for 20-40 seconds at 200mA (for a minigel). Then turn the current off and remove the liquid from the well.
6. Repeat the elution until the majority of the DNA band has been eluted. Pool the collected liquids.
7. Extract twice with phenol/chloroform/IAA.
8. Add 1/10 volume of 3M sodium acetate and 2 volumes of ethanol. Place at -70°C for 30 minutes.
9. Centrifuge for 15 minutes at 4°C, in a microfuge. Wash the pellet once with 70% ethanol and resuspend in a final volume of 10-20 µl of STE.

This method of fragment purification is preferable if the fragment to be purified is well separated from the other bands in the gel. The amount of DNA to load onto the gel should be empirically determined.

2.8.1.2 Using Dialysis Tubing

1. The gel should be run in the same way as before (Section 2.8.1.1).
2. Visualise the required DNA band, either by transillumination or with the hand held u.v. lamp.
3. Cut the band out with a scalpel and place the gel slice in a short piece of dialysis tubing (previously boiled in STE). Add 100-200 µl of running buffer and clip-seal the tubing.
4. Submerge the dialysis tubing in an electrophoresis tank. Turn the current on.
5. Run for between 30 minutes and 2 hours at 100 mA and then reverse the current for 30 seconds to bring the DNA off the dialysis tubing.
6. Remove the liquid from the dialysis tubing and treat in the same was as fragments that are collected in glycerol (Section 2.8.1.1, steps 7-9).

This method is preferable when the DNA fragment of interest is close to the other bands, for example, if a partial restriction digestion has been used to generate the DNA fragments.

2.8.2 DNA Fragments from Polyacrylamide Gels

1. Run and stain the gel as usual.
2. Visualise the DNA bands by transillumination.
3. Cut out the DNA band, as a gel slice, with a scalpel.
4. Transfer the gel slice to an eppendorf tube and grind to a

paste with a siliconised glass rod. Add 300 µl of elution buffer, mix, and incubate at 37°C, overnight, with shaking.

5. Fill the bottom of a 1 ml syringe with baked glass wool (approximately to the 0.1 ml mark). Transfer the contents of the eppendorf to the syringe and centrifuge at 500 rpm for 2 minutes at RT. Collect the liquid in an eppendorf.

6. Add a further 100 µl of elution buffer to the syringe and centrifuge again. Approximately 400 µl of liquid should be obtained.

7. Add 1.0 ml of ethanol and place at -70°C for 30 minutes.

8. Centrifuge for 15 minutes at 4°C in a microfuge. Resuspend the pellet in 200 µl of STE for 30 minutes at RT.

9. Add 10 µl of 3M sodium acetate and 500 µl of ethanol. Place at -70°C for 30 minutes.

10. Centrifuge for 15 minutes at 4°C. Wash the pellet with 100% ethanol and resuspend in 10 µl of STE.

Elution buffer:	ammonium acetate	500.0 mM
	magnesium acetate	10.0 mM
	SDS	0.1 %
	EDTA (pH8.0)	0.1 mM

2.9 DNA Manipulations

For more general information refer to Maniatis et al. (1982).

2.9.1 Restriction Digestion

Enzymes were purchased from either BRL or NBL and used at the specified concentrations. The following buffers were used

for restriction.

10X Low salt buffer: HaeIII, KpnI, ClaI, TaqI (at 65°C), RsaI.

NaCl	60 mM
Tris-HCl (pH7.5)	60 mM
$MgCl_2$	66 mM

10X Medium salt buffer: HindIII, AccI, BglII, HincII, AluI, PstI, PvuII, EcoRI, Sau3A.

NaCl	600 mM
Tris-HCl (pH7.5)	70 mM
$MgCl_2$	70 mM

10X High salt buffer: BamHI, SalI, XbaI, XhoI.

NaCl	1.5 M
Tris-HCl (pH7.5)	60.0 mM
$MgCl_2$	60.0 mM

10X "Universal" buffer: SmaI, NruI, HpaI and multiple digestions with enzymes that have different salt requirements.

Tris-base	1.00 g
potassium acetate	1.62 g
magnesium acetate (.4 H_2O)	0.54 g
spermidine	0.26 g
H_2O (final volume)	25.00 ml

adjust pH to 7.9 with glacial acetic acid add 0.02 g DTT and sterilise by filtration.

DNA that is obtained by the miniprep procedure is contaminated with large amounts of RNA. Therefore, restricted miniprep DNA needs to be treated with 1 µl of RNAase (1 mg/ml in 10 mM sodium acetate (pH5.0) and heated at 85°C for 10 minutes to denature any contaminating DNAase) for 5 minutes before being visualised on gels.

2.9.2 Dephosphorylation

The 5' terminal phosphate was removed from DNA to prevent self-ligation.

1. Digest the DNA to completion.
2. Add calf intestinal alkaline phosphatase (CIAP) (Boehringer) to the digestion. Use 1-2 U of enzyme per µg of DNA. This enzyme will work in any of the restriction buffers in Section 2.9.1.
3. Incubate at 37°C for 15 minutes.
4. Extract twice with phenol/chloroform/IAA.
5. Ethanol precipitate and resuspend the DNA as normal.

2.9.3 Fill-in Reaction

The recessed ends of double-stranded DNA were filled in to produce blunt ends. Make up the following reaction:

1. 10X polymerase buffer — 2 µl
   10X DTT — 2 µl
   Klenow fragment (BRL) — 1 U/µg of DNA
   DNA — x/µl

   H_2O (final volume) — 18 µl
2. Leave at RT for 5 minutes.
3. Add 2 µl of 10X dNTP mix.
4. Leave at RT for 30 minutes.
5. Add 30 µl of STE, phenol extract, ethanol precipitate and resuspend in STE.

10X polymerase buffer:	Tris-HCl (pH7.2)	500 mM
	$MgSO_4$	100 mM
	DTT	1 mM
10X dNTP mix:	dATP	1 mM
	dTTP	1 mM
	dCTP	1 mM
	dGTP	1 mM

2.9.4 Ligation of Sticky Ends

The efficiency of ligation is increased by the phenol extraction of the DNA.

1. Make up the following reaction:

10X ligation buffer	2.0 µl
DNA	x µl
H_2O (final volume)	17.5 µl

2. Heat at 70°C for 10 minutes.

3. Transfer to ice and leave for 5 minutes.

4. Add 2 µl of 10X SAD and 0.5 µl of T4 DNA ligase (BRL).

5. Incubate for 12-16 hours at 12°C.

6. Heat inactivate the ligase by incubation at 70°C for 10 minutes.

7. Use 5-10 µl for transformation into E. coli. The efficiency of transformation can be increased by the phenol extraction and ethanol precipitation of the ligation.

10X ligation buffer:	Tris-HCl (pH7.5)	600 mM
	$MgCl_2$	100 mM
	EDTA (pH8.0)	10 mM
10X SAD:	spermidine	10 mM
	ATP	20 mM
	DTT	100 mM

2.9.5 Ligation of Blunt Ends

1. Make up the following reaction mixture:

10X ligation buffer	1 µl
0.25% nonidet P-40	1 µl
DNA	x µl
H_2O (final volume)	7 µl

2. Heat and chill as for a sticky end ligation.

3. Add 1 µl of 10X SAD and 2µl of T4 DNA ligase.

4. Incubate for 6 hours at RT.

5. Heat inactivate and transform E. coli as before.

The DNA from a fill-in reaction can be ligated directly (for example, with synthetic linkers) without phenol extraction and ethanol precipitation. Blunt end ligation will proceed in the 10X polymerase buffer described in Section 2.9.3.

2.9.6 Bal31 Digestion

This exonuclease was used for the generation of deletion series.

1. Linearise the plasmid DNA with the appropriate restriction enzyme. Phenol extract, ethanol precipitate and resuspend in H_2O.

2. Make up a reaction mixture using the following ratios:

DNA	1.0 µg
5X Bal31 buffer	8.0 µl
H_2O (final volume)	38.8 µl

3. Pre-warm this mixture to 37°C.

4. Add 1.2 µl Bal31 (BRL, 1.16 U/µl), mix rapidly and return to

37°C.

5. Remove DNA samples at the desired time points and plunge into cold phenol/chloroform/IAA (on ice).

6. Extract, ethanol precipitate and resuspend in STE.

7. A fraction of each sample should be digested with a suitable restriction enzyme and visualised on a gel to determine the extent of exonuclease digestion.

8. The appropriate samples should be filled-in, blunt end ligated (with synthetic oligonucleotide linkers if required) and transformed into *E. coli*.

The rate at which any single preparation of Bal31 digests DNA should be empirically determined. For example, using the above conditions, Bal31 digested the *TYA* region of Tyl-15 in the 3' to 5' direction at a rate of 100-150 bp per minute.

5X Bal31 buffer	$CaCl_2$	60 mM
	$MgCl_2$	60 mM
	NaCl	3 M
	Tris-HCl (pH8.0)	100 mM
	EDTA (pH8.0)	5 mM

2.10 Determination of Nucleotide Sequences (Sanger *et al.*, 1977)

All the procedures are described in the M13 Cloning/"dideoxy" Sequencing Manual (BRL). DNA fragments, to be sequenced were cloned into the vectors M13mp18 or M13mp19 (Norrander *et al.*, 1983). All reagents were obtained from BRL.

2.11 *E. coli* Transformation

2.11.1 AKEC28

The following procedure reproducibly gave 10^6-10^7 transformants per μg of supercoiled plasmid DNA.

1. Streak *E. coli* cells to single colonies on a plate of LB and incubate at 37°C, overnight.
2. Transfer a single colony to 2 ml of LB (liquid media) and incubate at 37°C, overnight without shaking.
3. Inoculate 50 ml of LB with 1 ml of the overnight culture. Incubate at 37°C with good aeration.
4. Grow the cells for approximately 105 minutes, the OD_{600} should be 0.15-0.25.
5. Transfer the culture to a falcon tube and chill on ice for 10 minutes.
6. Centrifuge at 3 krpm for 5 minutes at 4°C and resuspend the cells in 25 ml of 50 mM $CaCl_2$, 50 μg/ml thymidine (Tdr).
7. Leave on ice for 5-20 minutes, harvest the cells and resuspend in 3 ml of $CaCl_2$/Tdr.
8. Leave on ice for between 10 minutes and 24 hours (I found that maximum competence was reached after approximately 6 hours).
9. Add 1-10 μl of DNA to 200 μl of the competent cells in an eppendorf. Leave on ice for 10 minutes.
10. Heat shock the cells at 37°C for 5 minutes.
11. Chill the cells on ice for 5 minutes.
12. Transfer the cells to 3 ml of prewarmed LB in a plastic tube and incubate for 60 minutes with shaking.

13. Harvest the cells by centrifuging at 3 krpm for 5 minutes at RT.

14. Decant the media and spread the cells on selective plates. Dilute the cells in sterile H_2O if necessary: the optimum number of colonies is 100-1000 per plate.

15. Incubate at 37°C overnight.

2.11.2 HB101

Using competent HB101 cells (BRL) made by a modification of the method of Hanahan (1983), a transformation frequency in excess of 10^8 transformants per μg of supercoiled plasmid DNA was obtained. The method is described by BRL (Cat. No. 8260SA).

2.12 Yeast Transformation (Hinnen et al., 1978)

Yeast cells were transformed by preparing spheroplasts with the enzyme glusulase.

Regeneration agar:

sorbitol	45.5 g
YNB, Difco	1.7 g
glucose	5.0 g
agar	7.5 g
H_2O (final volume)	250 ml

add tryptophan, leucine, uracil and histidine as required (for MD40-4c).

1. Streak a plate of yeast to single cononies on YPD (growth at 30°C).

2. Inoculate 2 ml of YPD with a single colony and grow overnight at 30°C.

3. Inoculate 100 ml of YPD (fresh) with an appropriate volume of the liquid preculture.

4. Incubate at 30°C with vigorous aeration (usually overnight).

5. Harvest the cells at a density of 1-2 x 10^7 cells per ml, by centrifugation at 3 krpm for 5 minutes at RT (use falcon tubes).

6. Wash the cells in 20 ml of 1 M sorbitol and resuspend in 10 ml of 1 M sorbitol.

7. Add 0.2 ml of glusulase (KH630, Endo Laboratories, Japan).

8. Incubate at 30°C for 1-2 hours, until spheroplasts are formed. Assay spheroplasting by adding a small amount of the cells to 5% SDS; the liquid should go clear as "ghosts" are formed.

9. Wash the spheroplasts twice with 1 M sorbitol and once with STC. Resuspend the spheroplasts in a final volume of 1 ml of STC.

10. Add 1-5 µg of plasmid DNA to 100 µl of the spheroplasts, in an eppendorf tube.

11. Leave at RT for 10-20 minutes.

12. Add 1.0 ml of PEG to the spheroplasts and mix well by inverting the eppendorf. Do not vortex. Leave at RT for 10 minutes (no longer).

13. Harvest the spheroplasts by centrifuging in a microfuge at RT. Do not allow the microfuge to reach full speed.

14. Pour off the liquid and resuspend the spheroplasts in 1.2 ml of 1 M sorbitol. Dilute the spheroplasts in 1 M sorbitol, if required.

15. Transfer the spheroplasts to 18 ml of selective liquid regeneration agar (45-50°C) and pour into petri dishes immediately.

16. Incubate for 3-5 days at 30°C or until colonies appear.

STC:	sorbitol	1 M
	$CaCl_2$	10 mM
	Tris-HCl (pH7.5)	10 mM

PEG: 44% PEG 4000 in STC.

All yeast strains that are transformed with a plasmid pX are designated TX.

2.13 Preparation of Proteins from Yeast

This method is a crude preparation of the cytoplasmic, soluble proteins.

1. Grow a yeast culture to the relevant density. If the proteins of an exponentially growing culture are being studied, the final cell density should be 4-8 x 10^6 cells/ml.

2. Harvest the cells and wash twice with 25 mM sodium phosphate (pH7.5).

3. Resuspend the cell pellet in 300 µl of cold 25 mM sodium phosphate, 1 mM PMSF. The PMSF is a protease inhibitor. The ideal pellet size is 3-4 mm up the side of a 1.5 ml eppendorf.

4. Add glass beads (40-mesh; BDH) so that the level of the beads is 2 mm below the meniscus of the cell suspension.

5. Vortex the mixture at 4°C for 1-2 minutes and keep on ice whilst not vortexing.

6. Remove the supernatant and wash the beads with a further 200 µl of 25 mM sodium phosphate/PMSF.

7. Pool the two supernatants (500 µl) and centrifuge for 15 minutes at 4°C.

8. Keep the supernatant and store at -20°C.

9. The concentration of proteins in the extracts can be measured by a dye-binding assay (Bradford, 1976) obtained from Bio-Rad Laboratories.

10. For analysis on SDS-polyacrylamide gels, 30 µg of proteins were loaded for Coomassie blue staining and 3 µg were loaded for Western blotting.

2.14 Preparation of Virus-Like Particles (VLPs) from Yeast (Mellor et al., 1985c)

1. Grow 1 litre of cells selectively at 30°C, with aeration to a density of 1 x 10^7 cells/ml.

2. Harvest the cells at 5 krpm for 10 minutes at RT (JA-10 rotor).

3. Wash the cells once with H_2O and resuspend in 1.0 ml of TEN, in a 15 ml corex tube.

4. Add glass beads to about half the volume of the cell suspension.

5. Vortex vigorously for 5 minutes at 4°C (the cell breakage should be greater than 70%).

6. Centrifuge at 3 krpm for 15 seconds at 4°C and remove the supernatant. Wash the glass beads with 2 ml of TEN and add this to the first supernatant.

7. Sonicate the pooled supernatants 10 times on ice for 15

second periods (6-7 μm peak-to-peak).

8. Centrifuge the crude extract at 20.7 krpm (40,000 g) for one hour at 4°C (SW50 rotor), to remove the cell debris.

9. Remove the supernatant (cleared extract) and layer onto 15-45% (w/v) sucrose gradient in 10 mM Tris-HCl (pH7.2), 10 mM NaCl. Centrifuge at 24 krpm (763,000 g) for 3 hours at 15°C (SW27 rotor). Prepare the sucrose gradients by layering 8 ml of 45, 35, 25 and 15% sucrose (BRL, enzyme grade) solutions into an SW27 polyallomer tube (Beckman) and leaving at 4°C overnight to equilibriate.

10. Fractionate the gradient into Nunc Cryotubes from the bottom of the gradient. These should be transferred to 4°C as soon as possible.

11. Assay the collected fractions for the position of the VLP peak. This can be done by reverse transcriptase assays, RNA analysis or protein analysis (50 μl aliquots of fractions if using SDS-polyacrylamide gels).

12. Once the fractions that contain VLPs have been determined, the VLPs can be pelleted by centrifugation at 32.7 krpm (100,000 g) for one hour at 4°C (SW50 rotor).

13. Resuspend the pellets in 100 μl of 10 mM Tris-HCl (pH7.2), 10 mM NaCl by shaking at 4°C, overnight.

TEN buffer:	Tris-HCl (pH7.2)	10 mM
	EDTA (pH.8.0)	2 mM
	NaCl	140 mM

This early method of VLP purification was later modified in the following ways by Adams et al. (1987).

1. Leave the washed cells (in the corex tube) at -70°C for one hour prior to breakage. This helps to break the cells.

2. Add 1 mM PMSF to the TEN buffer.

3. Omit the sonication step.

4. Wash the glass beads twice with 2 ml of TEN/PMSF.

5. Centrifuge the pooled supernatants for 15 minutes at 4°C in a microfuge, rather than with an SW50 rotor.

6. Centrifuge the cleared extract at 32.7 krpm (100,000 g) for one hour at 4°C, to pellet the VLPs. Resuspend the pellet in 0.5 ml of TEN/PMSF overnight at 4°C, centrifuge for 15 minutes 4°C in a microfuge, prior to loading onto the sucrose gradient.

2.15 Assay for Reverse Transcriptase Activity

The most commonly used assay was similar to that of Goff et al. (1981), using endogenous substrates.

1. Add 10 µl of sample (usually a sucrose gradient fraction) to 20 µl of an assay cocktail to give the following reaction concentrations:

Tris-HCl (pH8.3)	50 mM
DTT	20 mM
$MnSO_4$	0.6 mM
NaCl	60 mM
NP-40	0.05 %
dATP, dCTP, dGTP	0.5 mM
RNasin (P & S Biochemicals)	0.06 U/ml
(α-^{32}P)dTTP(400Ci/mmol, Amersham)	0.02 µCi/µl

2. Incubate at 30°C for 2 hours.

3. Transfer 15 µl of the reacton mixture to a DE81 filter (Whatman).

4. Wash the filters 3 times in 5% Na_2HPO_4 for 15 minutes, once in H_2O and twice in ethanol. Allow to air dry.

5. Count the dried filters in a scintillation counter.

Occasionally exogenous substrates were added. The reaction conditions were the same except that Ø.5 mg/ml of poly(rA) (Calbiochem) and Ø.Ø4 mg/ml of oligo (dT) (Collaborative Research) or Ø.5 mg/ml of poly (rC) (Calbiochem) and Ø.Ø4 mg/ml of oligo (dG) (Collaborative Research) or Ø.1 mg/ml of poly (rCm)-oligo (dG) (1:1 molar nucleotide ratio) (P-L Biochemicals) were present in each assay instead of the cold dNTPs. (α-^{32}P) dGTP (Amersham) was used instead of (α-^{32}P)dTTP in the G-C reactions.

2.16 Western Blot Analysis

2.16.1 SDS-Polyacrylamide Gel Electrophoresis (Laemmli, 1970)

The following types of gels were used. The acrylamide: bis-acrylamide ratios are shown in brackets.

Main gel:	9% (30:1.6)	12.5% (30:Ø.8)
~30% acrylamide	9.Ø ml	12.5 ml
1.5 M Tris-HCl (pH8.8)	7.5 ml	7.5 ml
H_2O	13.Ø ml	9.5 ml
10% ammonium persulphate	Ø.15ml	Ø.15ml
10% SDS	Ø.3 ml	Ø.3 ml
TEMED	20.Ø µl	20.Ø µl
Stacking gel (5%):		
~30% acrylamide	2.5 ml	2.5 ml
Ø.5 M Tris-HCl (pH6.8)	3.75ml	3.75ml
H_2O	8.4 ml	8.4 ml
10% ammonium persulphate	Ø.15ml	Ø.15ml
10% SDS	Ø.15ml	Ø.15ml
TEMED	20.Ø µl	20.Ø µl

1. Degas the acrylamide, Tris and H_2O for the main gel. Add the ammonium persulphate, SDS and TEMED.

2. Pour the main gel between the plates to within 3 cm of the top and overlay with 5% ethanol (in H_2O) and leave to set.

3. Make up the stacking gel in the same way.

4. Pour off the ethanol, rinse with H_2O and add the stacking gel. Clip in the well former and allow to set.

5. To prepare the protein samples for SDS-PAGE, add an equal volume of 2X Laemmli buffer to the sample and boil for 5 minutes prior to loading: this denatures the proteins. When analysing sucrose gradient fractions, add 50 µl aliquots to 15 µl of 4X Laemmli buffer (see Section 2.14, step 11).

6. Run the gel overnight. Typical running conditions for a 9% (30:1.6) gel are 80V for 16 hours.

2X Laemmli buffer:	Tris-HCl (pH6.8)	62.5 mM
	SDS	2.0 %
	glycerol	10.0 %
	BME	2.0 %
	urea	0.5 M
	bromophenol blue	2.0 mg/ml
	xylene cyanol FF	2.0 mg/ml
Running buffer:	glycine	14.4 g
	Tris-base	3.03g
	SDS	0.1 %
	H_2O (final volume)	1000 ml

All SDS-polyacrylamide gel analyses used prestained protein molecular weight standards (BRL) as size markers; the sizes (in kD) are 200 (myosin, H-chain), 97.4 (phosphorylase b), 68 (bovine serum albumin), 43 (ovalbumin), 25.7 (α-chymotrypsinogen), 18.4 (B-lactoglobulin) and 14.3 (lysozyme).

2.16.2 Coomassie Blue Staining

For the easy visualisation of proteins separated by SDS-PAGE, the gels were stained with Coomassie blue.

1. Stain the gel in 50% methanol, 10% acetic acid and less than 0.1% Coomassie blue for several hours (usually overnight).
2. Destain the gel in 20% methanol, 10% acetic acid for as long as necessary. The 30:1.6 gels never fully destain.
3. Store the gel in 10% methanol, 10% acetic acid in a sealed polythene bag.

2.16.3 Transfer of Proteins to Nitrocellulose (Towbin et al., 1979)

1. After electrophoresis, soak the gel in blotting buffer for 30 minutes, with one change.
2. Using a Bio-Rad Trans-blot apparatus, set up a sandwich composed of buffer-saturated "Scotch-Brite" pad, 2 MM paper, polyacrylamide gel, 0.45 μm nitrocellulose (BA85, Schleicher and Schuell), 3 MM paper and a second "Scotch-Brite" pad.
3. Insert the sandwich into the buffer filled apparatus (with a magnetic bar at the bottom, and a water coding system) so that the n/c faces the anode.
4. Run for 3 hours at 70V, by which time most of the protein should have been transferred.
5. The filter may be used immediately or stored indefinitely. To store the filter, rinse it with 20 mM Tris-HCl (pH7.5), 0.9% NaCl, dry at RT for 30 mins between 2 fresh pieces of 3MM paper,

wrap in Aluminium foil and place at -70°C.

Blotting buffer:	glycine	57.6 g
	Tris-base	12.12g
	methanol	800 ml
	H_2O (final volume)	4 litres

2.16.4 Hybridisation with Double Stranded DNA

1. After electro-blotting, wash the filter with 200 ml of the blocking, binding and washing buffer for 15 minutes at RT. Repeat for a further 15 minutes.

2. Transfer the filter to a polythene bag and add 1.0 ml of buffer and 100 µl of nick-translated total yeast DNA (see Section 12.18.5). The probe should not be denatured prior to hybridisation. Seal the bag and spread the hybridisation fluid evenly over the filter.

3. Leave for one hour at RT (protein side facing downwards).

4. Wash the filter four times in 200 ml of buffer for 30 minutes at RT.

5. Wrap the filter in clingfilm and expose to X-ray film (Kodak X-OMAT AR-5) at -70°C using Cronex (Dupont) lighting plus intensifying screens.

Blocking, binding and washing buffer:	NaCl	10.0mM
	Tris-HCl (pH7.9)	20.0mM
	PVP (polyvinyl pyrolidone)	0.1%
	ficoll	0.1%

2.16.5 Hybridisation with ^{125}I Labelled Monoclonal Antibody (Mellor et al., 1985a)

1. After electro-blotting, wash the filter twice in low-salt blocking buffer for 15 minutes at RT.
2. Transfer the filter to a polythene bag and add 1.Ø ml of the 2X hybridisation buffer plus 1.Ø ml of the monoclonal antibody (Mouse NK2 anti interferon antibody iodinated with ^{125}I, 12.6-19.5 μCi/μg, 7 x $1Ø^5$ dpm/ml). Seal the bag and incubate overnight at RT.
3. Wash the filter four times with low-salt washing buffer for 15 minutes at RT.
4. Expose to X-ray film at -7Ø°C.

Low-salt blocking buffer:	Tris-HCl (pH7.5)	2Ø.Ø mM
	NaCl	Ø.9 % (w/v)
	Triton X-1ØØ	Ø.Ø5% (v/v)
	casein (Hammersten BDH)	Ø.5 % (w/v)
2X hybridisation buffer:	Tris-HCl (pH7.5)	4Ø.Ø mM
	NaCl	1.8 % (w/v)
	Triton X-1ØØ	Ø.1 % (v/v)
	BSA (radioimmuno-assay grade)	1.Ø % (w/v)
Low-salt washing buffer:	Tris-HCl (pH7.5)	2Ø.Ø mM
	NaCl	Ø.9 % (w/v)
	Triton X-1ØØ	Ø.Ø5% (v/v)

2.16.6 Hybridisation with Non-Radiolabelled Antibody: Detection with an Immunoenzymatic Reaction (De Blas and Cherwinski, 1983)

1. Block the filter as in Section 2.16.5.
2. Incubate the filter in 1X hybridisation buffer (Section 2.16.5) with either neat serum (approximately a 1/1ØØ dilution)

or purifed IgG (Ø.1-Ø.5 µg/ml) overnight at 4°C in a polythene bag (as before).

3. Wash the filter three times with low-salt washing buffer (Section 2.16.5) for 2Ø minutes at RT.

4. Incubate the filter with the second antibody for one hour at RT. If the first antibody was raised in horse, a rabbit anti horse second antibody conjugated to horseradish peroxidase (Miles Scientific) can be used; the peroxidase reaction may then be developed (see step 7). If the first antibody was raised in rabbit, a goat anti rabbit second antibody (Sigma), at a 1/125 dilution, should be used.

5. In the case of the first antibody being raised in rabbit, wash the filter as before.

6. Incubate the filter with a 1/625 dilution of peroxidase-anti-peroxidase (PAP) complex raised in rabbit (Sigma) for one hour at RT.

7. Wash the filter as before.

8. Wash the filter for 1Ø minutes in 5Ø mM sodium phosphate (pH7.4) at RT.

9. Meanwhile, dissolve 5Ø mg of Diaminobenzidine tetrahydrochloride (DAB) (Sigma) in 1ØØ ml of 5Ø mM sodium phosphate (pH7.4), in a fume hood. Filter this through Whatman 3MM. Add, dropwise, 2-3ml of a 1% CoCl, 1% ammonium nickel sulphate solution followed by 6Ø µl of 3Ø% H_2O_2 (just prior to step 1Ø).

1Ø. Add the filter from step 8 to the solution prepared in step 9 and incubate for 5-1Ø minutes at RT.

11. Stop the reaction by removing the filter and washing with 50 mM sodium phosphate (pH7.4).
12. Dry the filter between 2 sheets of Whatman 3MM paper, wrap in clingfilm and store in the dark at RT.

2.16.7 Immunisation of Rabbits with VLPs

Standard immunisation procedures were used. Injections of 500 μg of protein were administered as an emulsion in Freund's complete adjuvant into the popliteal node region on day 1 and day 18 followed by a boost injection of 500 μg in incomplete adjuvant on day 64. All immunisations were carried out by Bob Sim (MRC Immunochemistry Unit, Oxford).

2.16.8 Preparation of Neat Serum from Blood

1. Chill the blood on ice for 2 hours in a falcon tube.
2. Centrifuge at 3 krpm for 20 minutes at $4^{o}C$.
3. Remove the supernatant, this is the serum.

2.16.9 Purification of IgG (Heide and Schwick, 1978)

1. Add 9 ml of H_2O to 1 ml of serum, in a siliconised 30 ml corex tube.
2. Add, dropwise, 10 ml of saturated ammonium sulphate.
3. Incubate at RT for 30 minutes.
4. Centrifuge at 5 krpm for 20 minutes at RT.
5. Resuspend the pellet in 2 ml of 0.5X PBS.

6. Dialyse for 4 hours at least 3 times against Ø.5X PBS at 4°C (total of 12 hours).

7. Add sodium azide to a final concentration of Ø.Ø1%.

8. Store at 4°C, the final concentration of IgG should be 5-1Ø mg/ml.

Saturated ammonium sulphate: leave overnight to ensure that it is saturated, remove the solution from the undissolved salt and pH to 7.Ø with ammonia.

1X PBS (pH7.4)	NaCl	8.Ø g
	KH_2PO_4	Ø.2 g
	Na_2HPO_4	1.15g
	KCl	Ø.2 g
	H_2O (final vol.)	1ØØØ ml

2.17 Electron Microscopy

All the electron microscopy data shown in this thesis was prepared at the University of Kent by Keith Gull. The procedures that were used will only be described briefly.

The thin sections (see Chapter 3) were prepared by the method of Byers and Goetsch (1975). The purified VLPs (see Chapters 3 and 6) were prepared by placing 5 ul of the resuspended pellet on a Formvar/carbon grid and allowing the VLPs to settle for a few minutes. A drop of 2% glutaraldehyde was added, followed by 2 or 3 drops of aqueous uranyl acetate. The sections were observed in a Phillips 41Ø transmission electron microscope.

2.18 Northern Blot Analysis

2.18.1 Preparation of Total Yeast RNA

1. Grow a yeast culture selectively at 30°C with vigorous aeration to a density of 4-6 x 10^6 cells/ml.
2. Harvest cells by centrifugation at 3 krpm for 5 minutes at 4°C.
3. Wash the cells twice with LET buffer. The final pellet should be 3-4 mm up the side of an eppendorf tube. This pellet may be stored indefinitely at -70°C.
4. Resuspend the cell pellet in 200 µl of cold LET buffer and add glass beads so that the level of the glass beads is 2 mm below the meniscus of the cell suspension.
5. Vortex hard for 20 seconds and return the tube to ice.
6. Add 100 µl of phenol (saturated with TEN buffer) and vortex for 45 seconds.
7. Return the tube to RT and add 100 µl of chloroform, 100 µl of H_2O (pH7.0) and 30 µl of 10X TNES (prewarmed to 55°C).
8. Vortex for 10 seconds and shake hard for 1 minute.
9. Centrifuge for 2 minutes at RT.
10. Remove the upper phase (contains the RNA) and extract with phenol/chloroform 3 times. By this time, there should be no interface between the phases.
11. Add 30 µl of 3M potassium acetate and 1.0 ml of ethanol. Mix well and place at -70°C for at least 30 minutes.
12. Centrifuge for 15 minutes at 4°C. There should be a large,

creamy white pellet.

13. Wash the pellet with 80% ethanol (20% water) and resuspend in 50 µl of water (pH7.0). Heat at 65°C for 5 minutes to remove any phenol induced aggregates. This pellet is mainly RNA, but contains a small amount of DNA (about 1%). Store at -70°C.

14. The concentration of RNA can be measured using the OD at 260 nm (OD_{260} 1= 40 µg of RNA). The yields vary from 50-150 µg.

LET buffer:	Tris-HCl (pH7.2)	100.0 mM
	LiCl	100.0 mM
	EDTA (pH8.5)	0.1 mM
10X TNES:	Tris-HCl (pH7.2)	100.0 mM
	NaCl	1.4 M
	EDTA (pH8.5)	10.0 mM
	SDS	1.0 %

2.18.2 Preparation of RNA from VLPs

1. Grow a 200 ml yeast culture selectively with vigorous aeration to a density of 6-8 x 10^6 cells/ml.

2. Harvest the cells and split into 2 halves. Wash one half twice with LET buffer and retain for total RNA extraction, wash the other half twice with TEN buffer.

3. Split the TEN washed cells into 2 eppendorfs and resuspend the cell pellets in 300 µl of TEN buffer. Add some glass beads.

4. Vortex hard for one minute at 4°C and return to ice.

5. Centrifuge for 2 minutes at 4°C to remove the glass beads and the cell debris from the cytoplasmic extract.

6. Carefully remove the supernatant and load onto a 15-45% (w/v) sucrose gradient in 10 mM Tris-HCl (pH7.2), 10 mM NaCl. Prepare the sucrose gradients as before (Section 2.14, step 9), except

use 2.5 ml steps in SW40 polyallomer tubes.

7. Centrifuge at 24 krpm for 3.5 hours at 15°C (SW40 rotor).

8. Collect 20 fractions of 0.5 ml from the bottom of the gradient, into Nunc tubes.

9. Add 200 µl of each fraction to 100 µl of phenol/chloroform, 25 µl of water (pH7.0) and 25 µl of 10X TNES.

10. Shake hard for 30 seconds and centrifuge for 2 minutes at RT.

11. Remove the upper phase and re-extract with phenol/chloroform.

12. Remove 200 µl of the upper phase and add 20 µl of 3M potassium acetate and 500 ul of ethanol. Mix well and place at -70°C for at least 30 minutes.

13. Centrifuge for 15 minutes at 4°C.

14. Wash the pellet once with 80% ethanol and resuspend in 20-30 µl of H_2O (pH7.0). Heat at 65°C for 5 minutes to remove phenol induced aggregates. Store at -70°C.

15. Run 10 µl of these RNAs on Northern gels. The RNA must be heated to 70°C for 5 minutes prior to aliquoting into the loading solution.

2.18.3 Running the Gel

1. Add 1.5 g of agarose (Sigma, No. A-6013 Type I: Low EEO) to approximately 150 ml of water. Heat in a microwave oven until the agarose has dissolved and the final volume is 102.4 ml.

2. Add 15 ml of 10X running buffer and 32.6 ml of formaldehyde (Fisons A.R. grade, 38% w/v). Pour the gel onto a baked glass plate; use greased-down, plastic side formers. The gel

dimensions should be 18 cm x 13 cm. Leave the gel to set for 1-2 hours.

3. Meanwhile, add 10 µl of each RNA sample to 8 µl of 10X running buffer, 14 µl of formaldehyde and 40 µl of deionised formamide. Heat at 55°C for 15 minutes and chill on ice for 5 minutes.

4. Add 8 µl of 10X gel loading buffer, mix and load onto gel.

5. Run the samples into the gel for 20-25 minutes at 100mA, or until the xylene cyanol FF has migrated 1.0 cm.

6. Submerge the gel with 1X running buffer and run for the desired amount of time (usually overnight). RNA that is 1.5 kb migrates close to the xylene cyanol FF.

10X running buffer (pH6.5):		
	1M NaH_2PO_4	64.6 ml
	1M Na_2HPO_4	35.4 ml
	0.4M EDTA (pH8.5)	25.0 ml
	H_2O (final volume)	1000 ml

Deionised formamide: deionise 500 ml of formamide (Fisons) by stirring with a magnetic bar with 3 g of Bio-Rad resin AG501-X8(D) (20-50 mesh) for 30 minutes at RT in the dark. If the OD_{270} is less than 0.2, separate the formamide from the resin by filtration through a Whatman filter (Glass microfibre paper, 2.4 cm GF/C): if the OD_{270} is greater than this, continue stirring. Divide the deionised formamide into 10 ml aliquots and store at -20°C in the dark.

10X gel loading buffer:		
	glycerol	50 %
	EDTA (pH8.5)	1 mM
	bromophenol blue	4 mg/ml
	xylene cyanol FF	4 mg/ml

2.18.4 Transfer of the RNA from the Gel to Nitrocellulose

All the RNA transfers in this thesis were unidirectional.

1. In a pyrex dish, invert half of a petri dish, put a baked glass plate (18 cm x 13 cm) on top and place two pieces of Whatman 3MM paper (soaked in 20X SSC) on top of the plate so that their edges touch the bottom of the pyrex dish. Fill the dish with 500-1000 ml of 20X SSC.

2. Invert the gel and place on top of the 3MM, carefully remove any bubbles from between the two. The bottom of the gel should be facing upwards.

3. Place strips of parafilm around the edges of the gel to prevent any buffer from bypassing the gel.

4. Place an 18 cm x 13 cm piece of 0.1 µm nitrocellulose (PH79, Schleicher and Schuell), on top of the gel. Ensure that there are no air bubbles between the gel and the filter. The filter should be boiled in water for 5 minutes and soaked in 20X SSC for 45 minutes prior to setting up the transfer.

5. Place 2 pieces of 3MM, soaked in 20X SSC, on top of the n/c, followed by 2 cm of dry 3MM and 3 cm of paper towels. On top of the stack, place a glass plate and seal the whole set-up with clingfilm. Put a weight (approx. 500 g) on top.

6. Leave to transfer for at least 18 hours: I left most transfers for 48 hours because better transfer of the 5.7 kb Ty RNA occurred with extended transfer: extra 20X SSC should be added after 24 hours, to avoid the gel cracking.

7. After transfer, allow the n/c to air dry and then bake at 80°C for 4 hours in a vacuum oven evacuated to 30 mm/Hg.

8. The n/c may be stored indefinitely in a sealed plastic bag.

20X SSC:		
	NaCl	175.32 g (3M)
	Tri-sodium citrate	85.20 g (Ø.3M)
	H_2O (final vol.)	1ØØØ ml

2.18.5 "Nick Translation" (Rigby et al., 1977)

1. To radiolabel DNA, use a nick translation kit (BRL) to set up the following reaction:

DNA fragment (Ø.1 ug)	1.Ø µl
PolI	1.Ø µl
dATP, dCTP, dGTP mix (Ø.2 mM)	1.Ø µl
(α-^{32}P)dTTP (Amersham)	Ø.5 µl
H_2O	6.5 µl

PolI: Ø.4 µg/µl DNA polymerase I, 4Ø pg/µl DNaseI, 5Ø mM Tris-HCl (pH7.5), 5 mM magnesium acetate, 1 mM BME, Ø.1 mM PMSF, 50% (v/v) glycerol and 1ØØ µg/ml nuclease-free BSA.

2. Incubate at 15°C for one hour.

3. Add 9Ø µl of STE and layer onto the top of a sephadex G-5Ø "spin column"; this separates the labelled DNA from the unincorporated (α-^{32}P)dTTP.

4. Centrifuge at 3.2 krpm for 2-4 minutes at RT. Collect the liquid from the bottom and store at -2Ø°C until required.

5. To calculate the specific activity of the probe, precipitate 2 µl of the final volume with 4 ml of 1Ø% tri-chloro acetic acid, use 1.Ø mg of denatured salmon sperm DNA as a carrier. Collect the precipitate on a GF/C filter and count in a scintillation

counter. The specific activity should be in excess of 10^7 cpm/µg DNA.

Sephadex G-50 "spin column" (Maniatis et al., 1982): to make this column, stuff "0.1" ml of glass wool into the bottom of a 1 ml syringe and fill with sephadex G-50 (medium) that has been swollen and equilibriated with STE. Centrifuge at 3.2 krpm for 2-4 minutes at RT. Place the column in a 15 ml corex tube with a 350 µl eppendorf tube at the bottom, it is now ready to use.

2.18.6 Hybridisation

1. For prehybridisation, soak the n/c on top of 200 ml of 10 mM Tris-HCl (pH7.5), in a baked pyrex dish, until thoroughly wet. Boil the n/c for 5 minutes.

2. Meanwhile make up the following hybridisation solution:

deionised formamide	10.0 ml
20X SSC	2.0 ml
5M sodium phosphate (pH6.5)	0.2 ml
10% SDS	0.2 ml
50X Denhardt's solution	2.0 ml
salmon sperm DNA (10 ug/ml)	0.5 ml
H_2O	5.0 ml
nick-translated probe	0.1 ml

The salmon sperm DNA, H_2O and probe should be boiled together for 5 minutes prior to adding the rest of the solutions.

3. Put the n/c in a polythene bag and add the hybridisation solution. Heat seal the bag to exclude all bubbles.

4. Incubate at 42°C with shaking for at least 18 hours. Have the side of the n/c, that has the RNA on it, facing downwards.

5. After hybridisation, wash the unhybridised probe off the n/c using 2X SSC, 50 mM sodium phosphate (pH6.5) and 0.1% SDS at

42°C. This should take approximately one hour.

6. Wrap the n/c in clingfilm and expose to X-ray film (see Section 2.16.4).

5M sodium phosphate (pH6.5):	Na_2HPO_4	44.Ø g
	NaH_2PO_4	1Ø7.6 g
	H_2O (final volume)	2ØØ.Ø ml

5ØX Denhardt's solution:	ficoll	Ø.5 g
	PVP (polyvinyl pyrolidone)	Ø.5 g
	BSA (pentax grade)	Ø.5 g
	H_2O (final volume)	5Ø.Ø ml

sterilise by filtration.

2.18.7 Quantitation

Bands were cut out of the n/c after autoradiography and counted in a scintillation counter. A previously exposed autoradiograph was used to align the bands with the n/c.

CHAPTER 3. THE IDENTIFICATION AND CHARACTERISATION OF TY VIRUS-LIKE PARTICLES

3.1. Introduction

From the discussion in Section 1.5, it is apparent that Ty elements strongly resemble integrated retroviral proviruses and eukaryotic retrotransposons such as copia and IAP genomes. All of these elements share similar genome organisations, consensus sequences (including that for a reverse transcriptase) and expression strategies. Ty has also been described as a retrotransposon by virtue of its genomic organisation and requirement for reverse transcriptase during its transposition via an RNA intermediate (Hauber et al., 1985; Clare and Farabaugh, 1985; Boeke et al., 1985). If the previously identified Ty proteins are analogous to those encoded by other retro-elements such as retroviruses, copia and IAP genomes, they might be expected to assemble into virus-like particles (VLPs) that contain Ty RNA and reverse transcriptase activity; such particles would be similar to the copia-VLPs of Drosophila (Shiba and Saigo, 1983) and the IAPs of certain rodent cell lines (Kuff et al., 1968).

3.2. The Identification of Ty Virus-Like Particles in Yeast.

Ty proteins have been identified previously by the massive overexpression of Ty ORFs using the powerful yeast expression vector pMA91 (Figure 1.8) (Mellor et al., 1983; Kingsman and Kingsman, 1983). The insertion of the whole transcriptional unit of Ty1-15 as a BamHI fragment (fragment 1Ø, Figure 1.8) into the BglII site of pMA91 yields a plasmid known as pMA91-1Ø (Dobson et

al., 1984): yeast transformants of pMA91-1Ø are designated T91-1Ø. The amount of 5.7 kb Ty RNA in T91-1Ø is 2Ø fold higher than that produced by the 3Ø-35 endogenous Ty elements. This increased level of Ty RNA led to the identification of six Ty proteins in T91-1Ø (Section 1.4.3, Dobson et al., 1984, Mellor et al., 1985b; Malim et al., 1987a). These Ty proteins have been designated p1 (5Ø kD), p2 (48 kD), p3 (19Ø kD), p4 (38 kD), p5 (33 kD) and p6 (28 kD).

If Ty proteins are able to form virus-like particles, such structures should be more abundant in cells that are overexpressing the Ty proteins (T91-1Ø) than in cells containing the vector, pMA91, alone (T91). Therefore, thin sections of the strains T91-1Ø and T91 were examined directly in the electron microscope. The cytoplasm of each T91-1Ø cell was densely packed with particles (Figure 3.1) whereas T91 cells contained relatively few particles (Figure 3.2). The particles were spherical, ~6Ø nm in diameter and covered 2Ø-8Ø% of the surface area of the all the T91-1Ø sections that were examined. These particles have been called Ty virus-like particles (Ty-VLPs) because of their much greater abundance in cells that were overexpressing the Ty proteins (Mellor et al., 1985c). The identification of Ty-VLPs reinforces the idea that Ty elements are closely related to copia elements, IAP genomes and retroviral proviruses. The sizes of retrotransposon VLPs are similar but their cellular locations are very different; copia-VLPs (5Ø nm) are nuclear and, although Ty-VLPs (6Ø nm) and IAPs (7Ø nm) are both cytoplasmic, IAPs are associated with the endoplasmic

Figure 3.1. Electron micrograph of a thin section of a T91-10 cell (X 7,600).

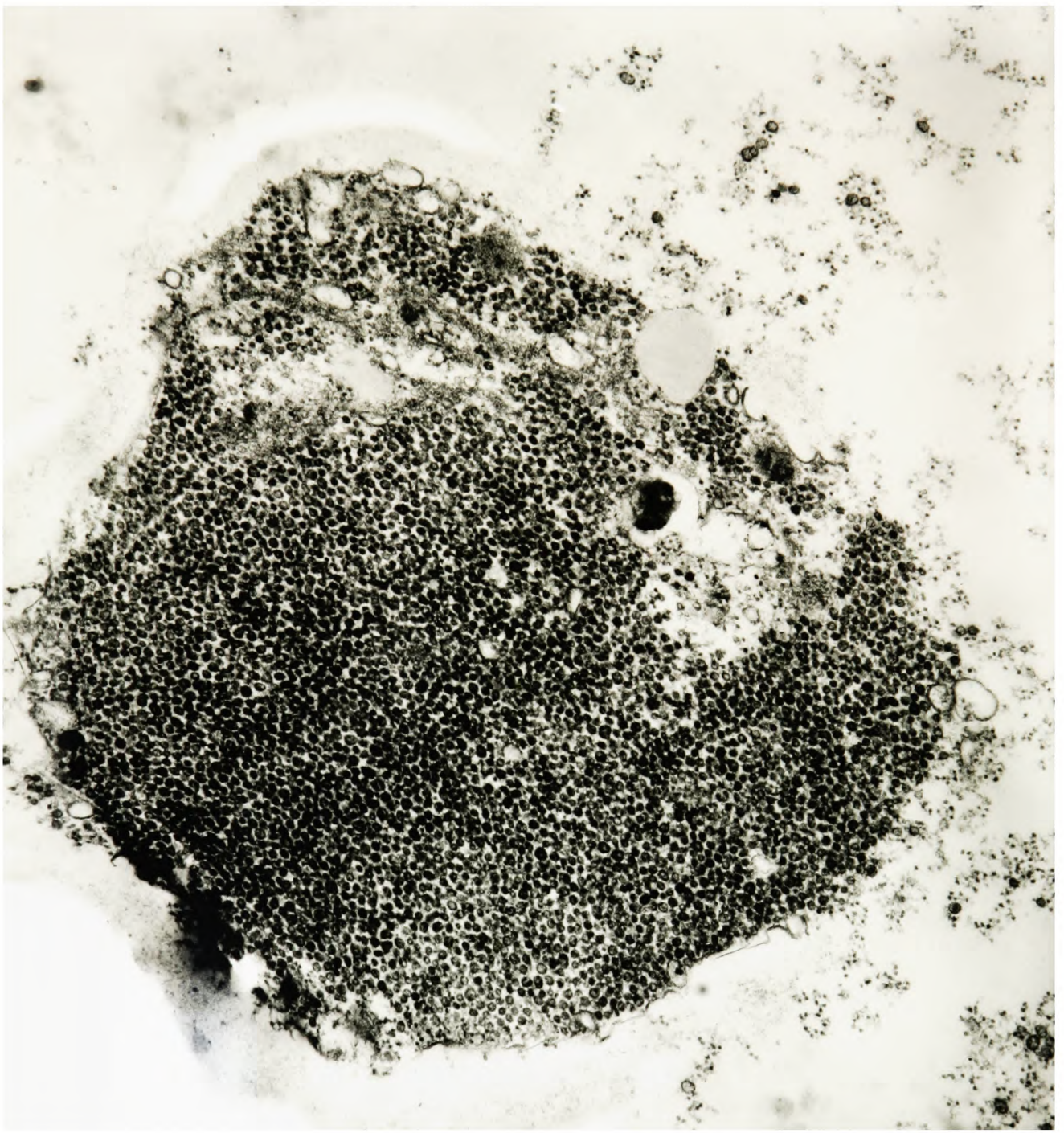
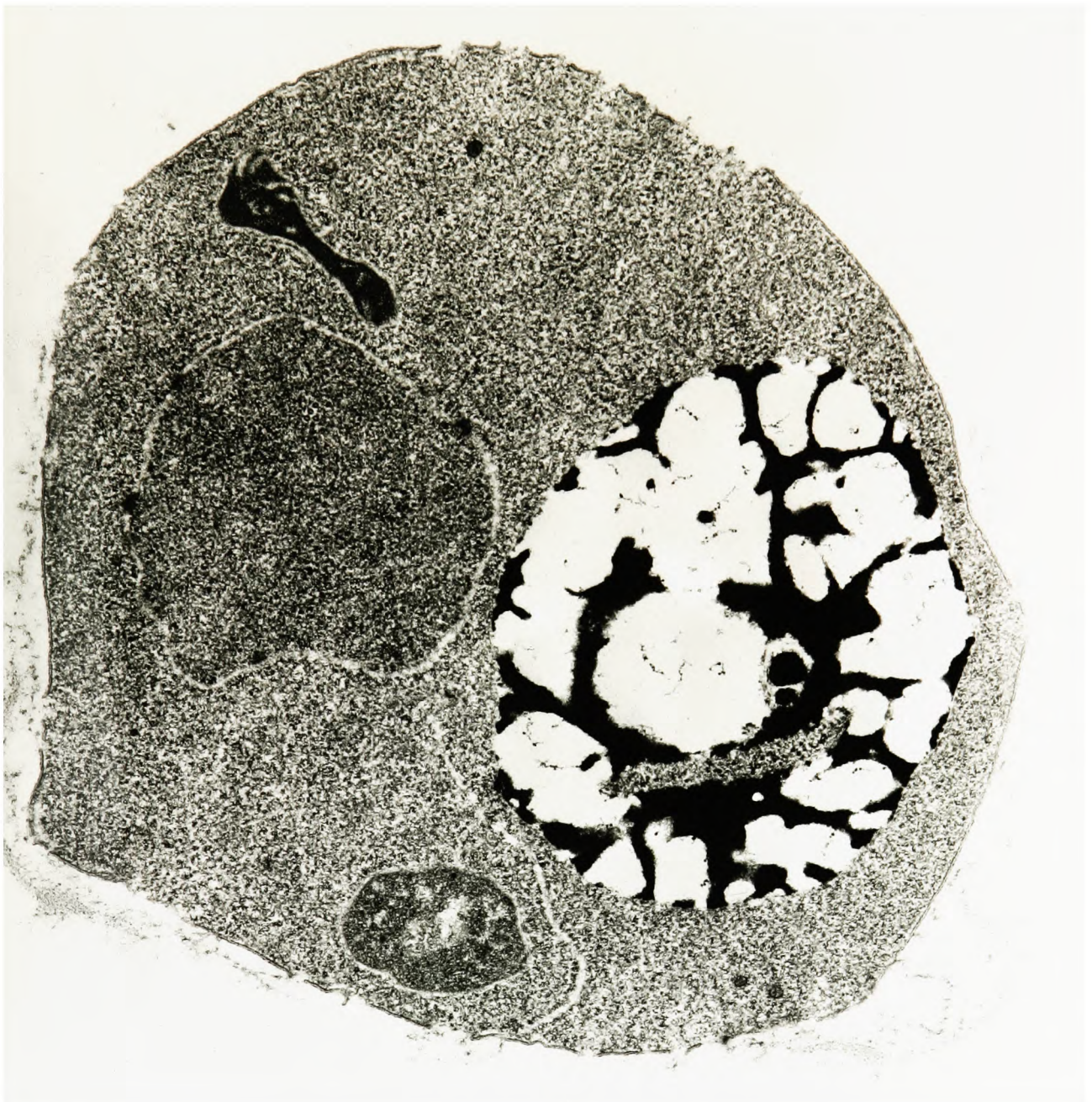


Figure 3.2. Electron micrograph of a thin section of a T91 cell (X 7,600).



reticulum whilst Ty-VLPs are apparently free from any membranes.

3.3. The Composition of Ty Virus-Like Particles

If the Ty-VLPs identified in Section 3.2. are analogous to the VLPs of other retrotransposons and retroviral particles, they ought to be composed of Ty encoded proteins (this has only been implied so far) and contain both RNA homologous to Ty and reverse transcriptase. To address these questions, an approach similar to that used by Shiba and Saigo (1983) to analyse copia-VLPs, was adopted. Cytoplasmic extracts from T91-1Ø and T91 were fractionated on SW27 (Section 2.14) and SW4Ø (Section 2.18.2) 15-45% (w/v) sucrose gradients. Each gradient was centrifuged for 3 hours at 15^{o}C. The SW27 gradient was collected in 3Ø fractions, each one of which was assayed for its endogenous reverse transcriptase activity; the assay used was based on that used by Goff *et al.* (1981) and it relied upon both endogenous template (Ty RNA) and endogenous primer (the hypothesised $tRNA^{met}$ binding to the 5.7 kb Ty RNA) being present in the VLPs. The SW4Ø gradient, however, was collected in 2Ø fractions, the RNA content of each was analysed by Northern blotting. The gradient profiles of reverse transcriptase activity and RNA content were aligned for the T91-1Ø extracts (Figure 3.3). There were clear coincident peaks of reverse transcriptase activity and 5.7 kb Ty RNA, corresponding to a particulate fraction with a sedimentation coefficient of ~175S. The T91 extracts displayed minor peaks of reverse transcriptase activity (Figure 3.3) and Ty RNA (data not shown) that can be attributed to the 3Ø-35 endogenous Ty

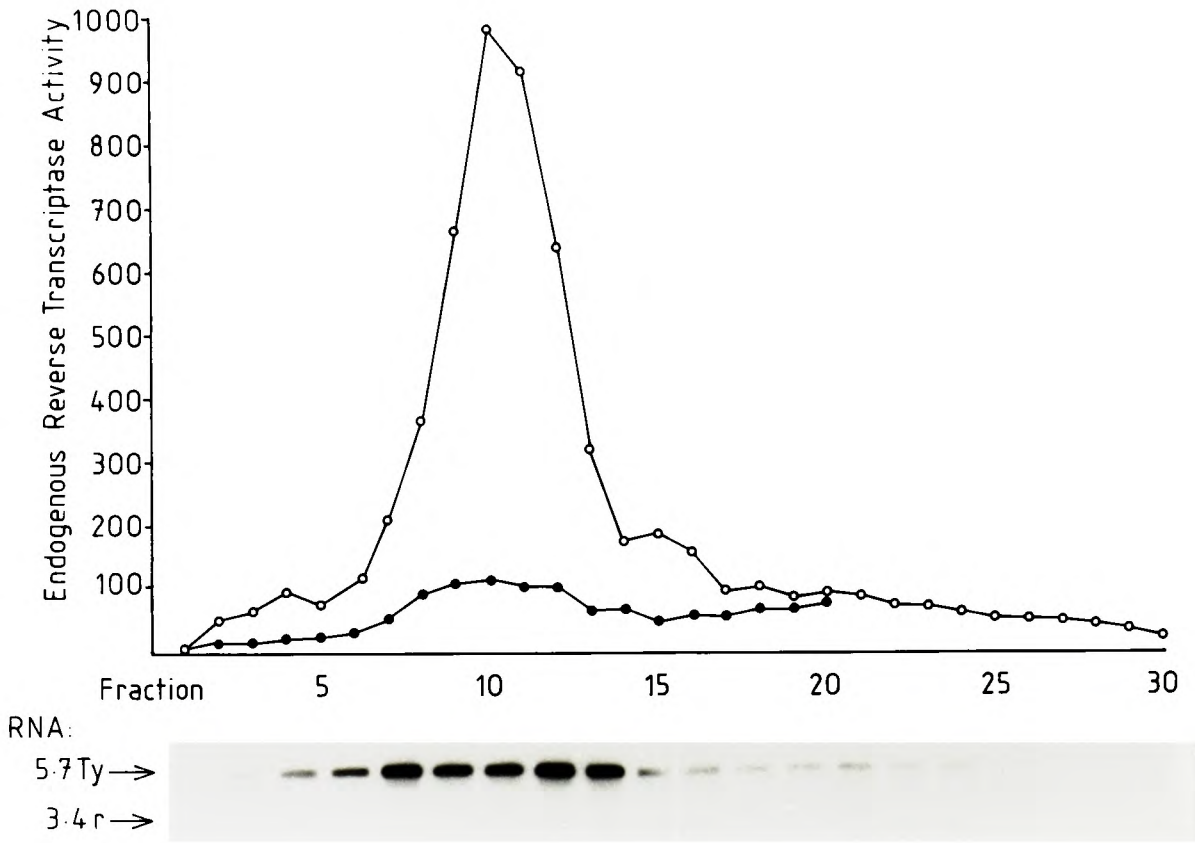
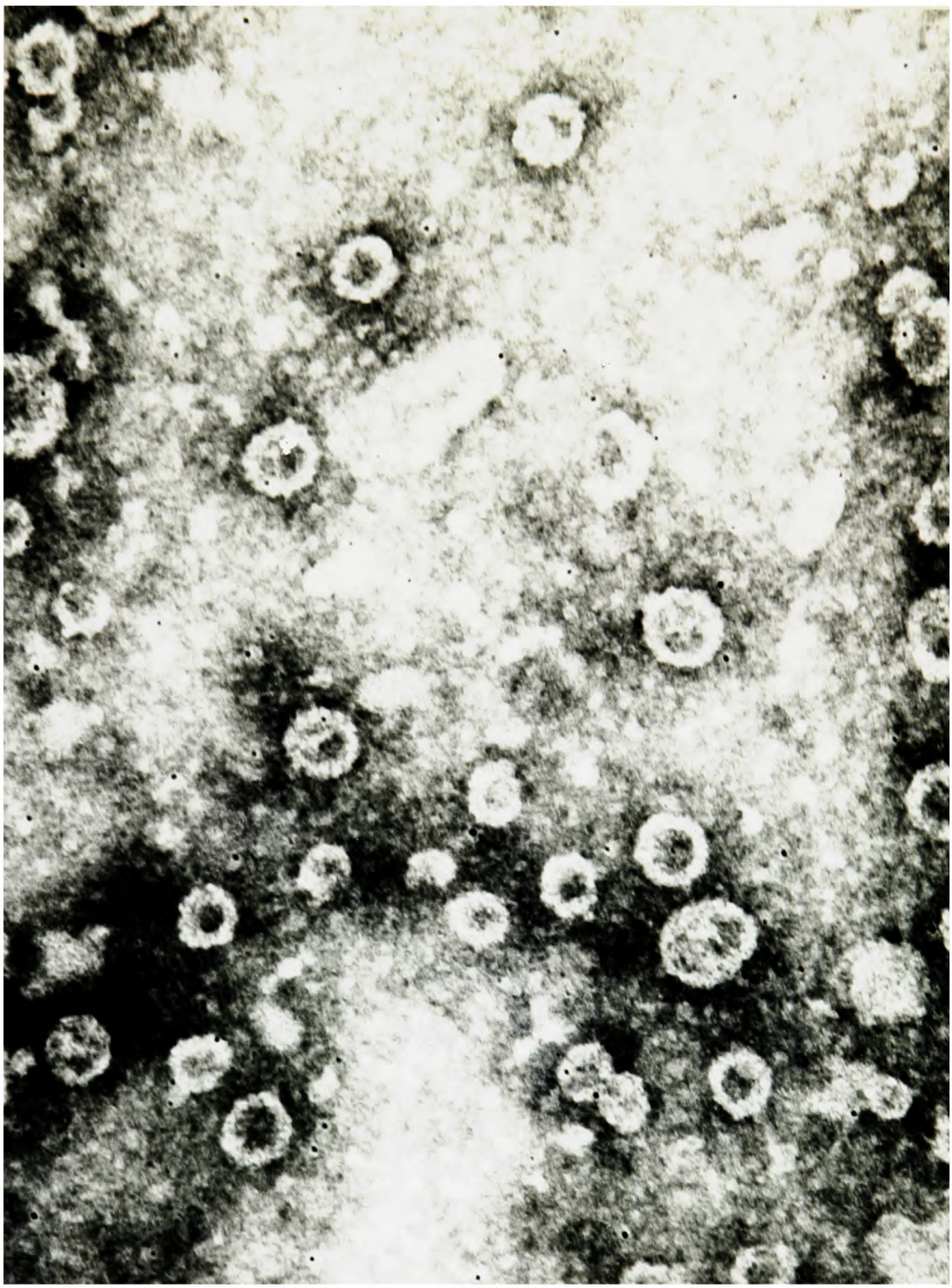


Figure 3.3. Sucrose gradient fractionation of extracts from T91-1Ø and T91 cells. The graph shows the endogenous reverse transcriptase activity (in c.p.m.) of fractions from both T91-1Ø (open circles) and T91 (solid circles) SW27 gradients. Below is an autoradiograph showing the Northern blot analysis of the RNA profile of a T91-1Ø SW4Ø gradient. The hybridisation probe was nick-translated 11 fragment DNA.

elements. The peak fractions from the T91-1Ø sucrose gradient (SW27, fractions 9-12 in Figure 3.3) were pooled, pelleted at 1ØØ,ØØØ g, resuspended and examined in the electron microscope (Figure 3.4). Ty-VLPs like those seen in the cross sections of T91-1Ø cells (Figure 3.1) were present at a high level of purity, the only visible contaminant being a few membrane fragments. The absence of 6Ø nm particles from the non-peak fractions (data not shown) suggests that these Ty-VLPs have associated reverse transcriptase activity and contain the 5.7 kb Ty RNA. The Ty-VLPs are relatively stable, they have a regular spherical shape and possess a bipartite shell with an electron dense core.

In an attempt to identify the proteins that form the structural component of Ty-VLPs, cytoplasmic extracts from T91-1Ø and T91 were fractionated on SW27 15-45% (w/v) sucrose gradients in the same way as before, except that only 2Ø fractions were collected instead of 3Ø (Figure 3.3). Two SDS-polyacrylamide gels were run and stained with Coomassie Blue, one was loaded with 5Ø µl aliquots of the fractions from the T91-1Ø gradient (Figure 3.5) whilst the other was similarily loaded with aliquots from the T91 gradient extract (Figure 3.6). There is a broad peak of proteins between fractions 3 and 9 of the T91-1Ø extract that were largely absent from the fractions corresponding to the top of the gradient (fractions 13-2Ø, Figure 3.5); a similar but rather faint peak of protein is visible between fractions 5 and 9 of the T91 extract (Figure 3.6). These proteins represent particles that sediment through sucrose gradients at a greater rate than the other yeast proteins present in T91-1Ø and T91

Figure 3.4. Electron micrograph of uranyl acetate-stained Ty-VLPs from the pelleted fractions of a SW27 sucrose gradient fractionated T91-1Ø extract (X 72,8ØØ).



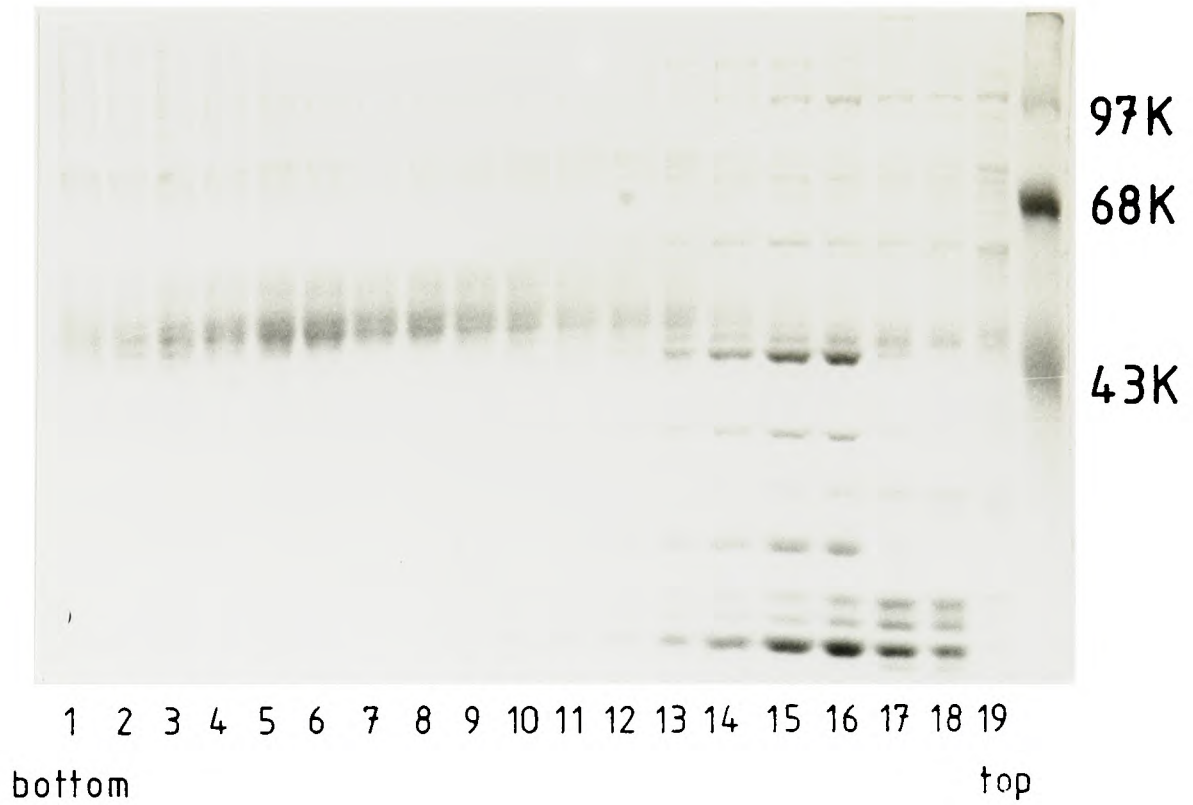


Figure 3.5. Coomassie Blue-stained gradient profile of a T91-10 extract. 50 µl aliquots of fractions from a SW27 15-45% (w/v) sucrose gradient treated T91-10 extract were denatured, loaded onto a 9% (30:1.6) SDS-polyacrylamide gel, electrophoresed and stained with Coomassie Blue.

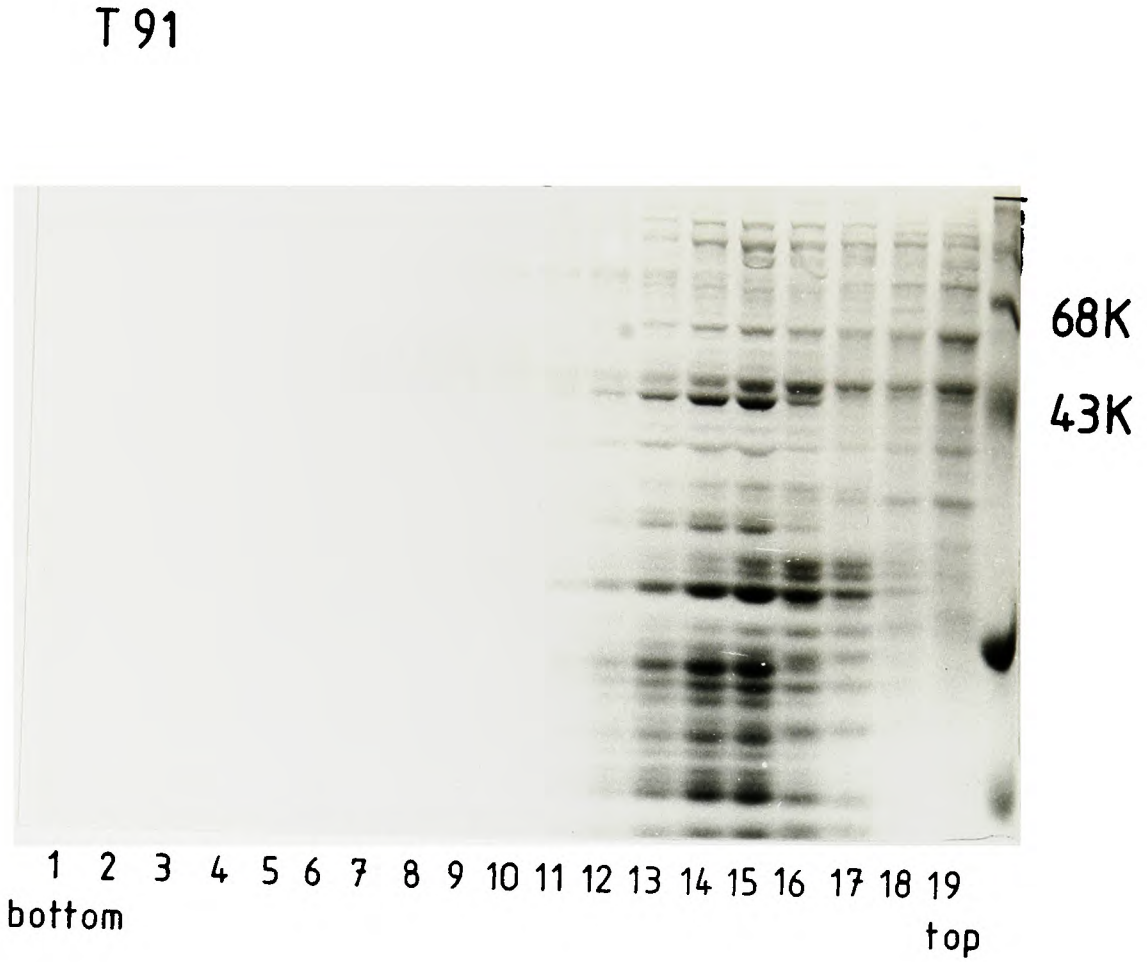


Figure 3.6. Coomassie Blue-stained gradient profile of a T91 extract. The preparation and treatment of the samples and the gel were described in the legend to Figure 3.5.

(these remain in the top fractions). The protein peak in Figure 3.5 coincides well with the peaks of reverse transcriptase activity and 5.7 kb Ty RNA that have been shown to be associated with Ty-VLPs, it is also much more intense than the peak in the T91 extract (contains relatively few Ty-VLPs). By deduction, these proteins are likely to be the proteins that form the structural components of Ty-VLPs. The most abundant protein in the T91-1Ø peak is ~48 kD, though it does appear as a doublet, whilst the minor band is ~5Ø kD (Figure 5.5). These sizes correspond to the previously identifed Ty proteins, p2 (48 kD) and p1 (5Ø kD) (Mellor et al., 1985b). There is evidence to suggest that the Ty protein p2 can, in fact, exist in two forms with slightly different gel mobilities (Adams et al., unpublished observation). The minor protein peak in T91 is also ~5Ø kD (Figure 3.6) and coincides with the small peak or reverse transcriptase activity that was detected in the fractionated T91 extract (Figure 3.3). This presumably corresponds to the low levels of Ty-VLPs that are produced by the 3Ø-35 endogenous Ty elements. It therefore seems that, like copia-VLPs and IAPs, the proteins that form Ty-VLPs are encoded by the Ty elements themselves.

It was described earlier (Section 1.4.5) that the Ty proteins are able to bind double stranded DNA in a crude binding assay (Mellor et al., 1985b). To test whether the major proteins of Ty-VLPs are indeed the Ty proteins p1 and p2, peak fractions from a T91-1Ø, SW27 sucrose gradient were pelleted and tested for their ability to bind dsDNA by probing a Western blot with nick-translated total yeast DNA (Section 2.16.4). Figure 3.7

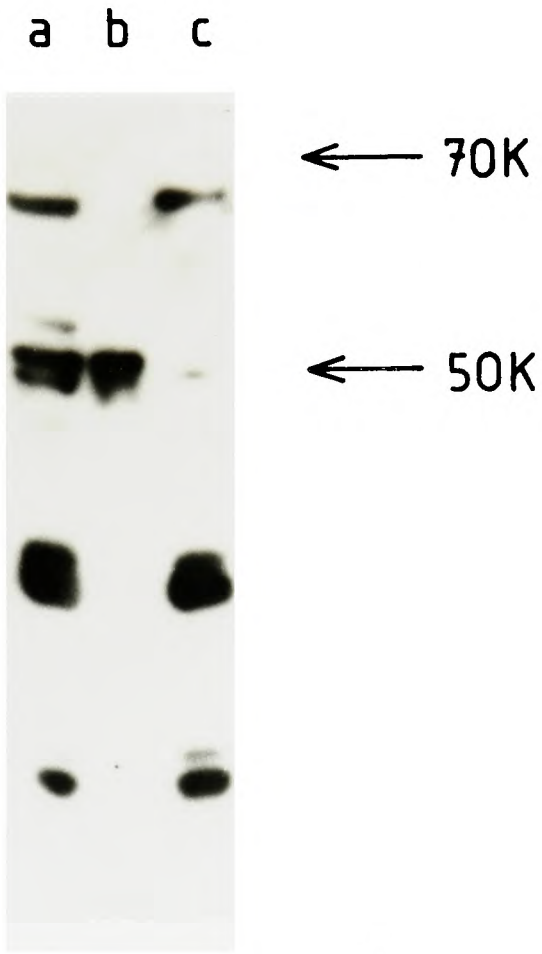


Figure 3.7. DNA-binding profiles of the separated proteins from a T91-1Ø extract (lane a), T91-1Ø Ty-VLPs (lane b) and a T91 extract (lane c). The proteins were separated on a 9% (3Ø:1.6) SDS-polyacrylamide gel and the filter was probed with nick-translated DNA isolated from MD4Ø-4c.

demonstrates that the Ty-VLP proteins do bind dsDNA (lane b) and that these proteins are the same ones that are present in T91-1Ø extracts (lane a, overexpressing Ty proteins) but virtually absent from the T91 extracts (lane c). This represents strong evidence that the major protein of Ty-VLPs is the Ty encoded protein p2, but that some p1 is also present.

3.4. The Properties of the Ty-VLP Associated Reverse Transcriptase

A series of experiments were carried out (Mellor et al., 1985c) to verify that the incorporation of radioactivity onto DE81 filters (Figure 3.3) was due to reverse transcriptase activity and not an alternative DNA polymerase, terminal deoxynucleotidyl transferase or nucleotide binding protein.

Firstly, the ^{32}P-labelled products of an endogenous reverse transcriptase assay (fraction 1Ø from the SW27 gradient in Figure 3.3 was used) were phenol extracted, ethanol precipitated and used in a dot-blot hybridisation experiment. Hybridisation occurred predominantly to Ty DNA, indicating that the products, and therefore the substrate of the reaction, both have Ty homology; the predicted substrate is the 5.7 kb Ty RNA that is associated to Ty-VLPs (Figure 3.3). Low level hybridisation to ribosomal DNA also occurred, indicating the slight contamination of the Ty-VLPs with ribosomal RNA (data not shown).

Secondly, the labelled products were run on a denaturing polyacrylamide gel and shown to consist of fragments a few

hundred nucleotides long (data not shown). These fragments were shown to be DNA by their sensitivity to DNaseI treatment and their non-appearance when the assay was carried out with ^{32}P-UTP rather than ^{32}P-dTTP (UTP can only be incorporated into RNA).

Thirdly, a detailed analysis of the enzymatic activity was carried out, the results of which are shown in Table 1. The endogenous activity was not inhibited by Actinomycin D, the presence of which inhibits DNA dependent DNA polymerases. Pretreatment of the Ty-VLPs with RNase A reduced the enzyme activity, indicating the utilisation of an RNA substrate; the partial protection afforded by the Ty-VLPs is reminiscent of copia-VLPs (Shiba and Saigo, 1983) and retroviruses (Baltimore, 1970). The requirement for a divalent cation (either Mg^{2+} or Mn^{2+}) is a characteristic of retroviral reverse transcriptases, as is its increased sensitivity to the assay in the presence of the non-ionic detergent nonidet P40. The Ty-VLP reverse transcriptase has a temperature optimum of approximately 20^{o}C, which probably explains why Ty transposition occurs most frequently at this temperature (Paquin and Williamson, 1984). The utilisation of the synthetic substrate-primers verified that the Ty-VLP associated activity was 'reverse transcriptase: poly(rC)-oligo(dG) is a poor substrate for DNA-dependent DNA polymerases (Chandra and Steel, 1977) whilst poly(rCm)-oligo(dG) is not a substrate at all (Gerard et al., 1974).

Table 1. The properties of the Ty-VLP reverse transcriptase

Assay	Variation	c.p.m. incorporated	Activity
Endogenous	Incubation at 20°C	2,249	100
	$-Mn^{2+}-Mg^{2+}$	50	2
	-NP40	1,086	48
	+ Actinomycin D	2,259	100
	+RNase A	547	24
	Incubation at 10°C	1,725	77
	Incubation at 30°C	1,927	86
	Incubation at 37°C	1,443	64
	Incubation at 42°C	1,041	46
poly(rC)-Oligo(dG)		4,188	
poly(rCm)-oligo(dG)		2,505	

3.5 Discussion and Subsequent Results

Ty virus-like particles (Ty-VLPs) have been identified in the cytoplasm of yeast cells by the massive overexpression of the whole transcriptional unit of a single Ty element. Ty-VLPs have a diameter of approximately 60 nm and are spherical with an electron dense core and a bipartite shell (Figure 3.4). They contain the genomic 5.7 kb Ty RNA and have associated reverse transcriptase activity: the major protein is the Ty encoded p2 (48 kD). Most of this work has been reported by Mellor et al. (1985c).

Ty-VLPs have also been identified by Garfinkel et al. (1985) and, furthermore, they have demonstrated that the particles are non-infectious since a marked Ty element could not be transferred from one yeast cell to another by co-cultivation. The 60 nm Ty-VLPs therefore resemble the 50 nm copia-VLPs and the 70 nm IAPs; none of these virus-like particles are infectious. This

may be because of the lack of an ORF corresponding to the *env* gene of retroviruses (Figure 1.4), the gene product of which forms the outer shell of retroviral particles. Its absence may also explain the smaller size of Ty-VLPs, copia-VLPs and IAPs compared to the 80-120 nm diameter of retroviral particles. It would be interesting to know whether the copia-like retrotransposon 17.6 (Figure 1.4) that possesses the coding capacity for an *env*-like protein, is able to form "infectious" VLPs.

As an extension of this work, Adams *et al*. (1987) have made a number of observations. Firstly, by raising an anti Ty-VLP antiserum in rabbits it has been shown that Ty-VLPs from T91-10 contain small amounts of the previously identified Ty proteins p4 (38 kD), p5 (33 kD) and p6 (28 kD) as well as p2 (48 kD) and p1 (50 kD). Secondly, the fractionation of a T91-11 extract, that only overexpresses the *TYA* region (Figure 1.8), through a sucrose gradient and subsequent analysis on an SDS-polyacrylamide gel (in the same way as Figure 3.5) shows that p1, the primary translation product of *TYA*, also forms particles (Figure 3.8). These particles were pelleted and examined in the electron microscope (Figure 3.9); they also have a diameter of ~60 nm but are less stable than the T91-10 Ty-VLPs, and have an electron-luscent core with a dark-staining outer shell. Thirdly, Cleveland mapping experiments showed that all the proteins in T91-10 Ty-VLPs share common amino acid sequences. Therefore it seems likely that the smaller proteins p2, p4, p5 and p6 are proteolytic cleavage products of p1. Fourthly, the enzyme

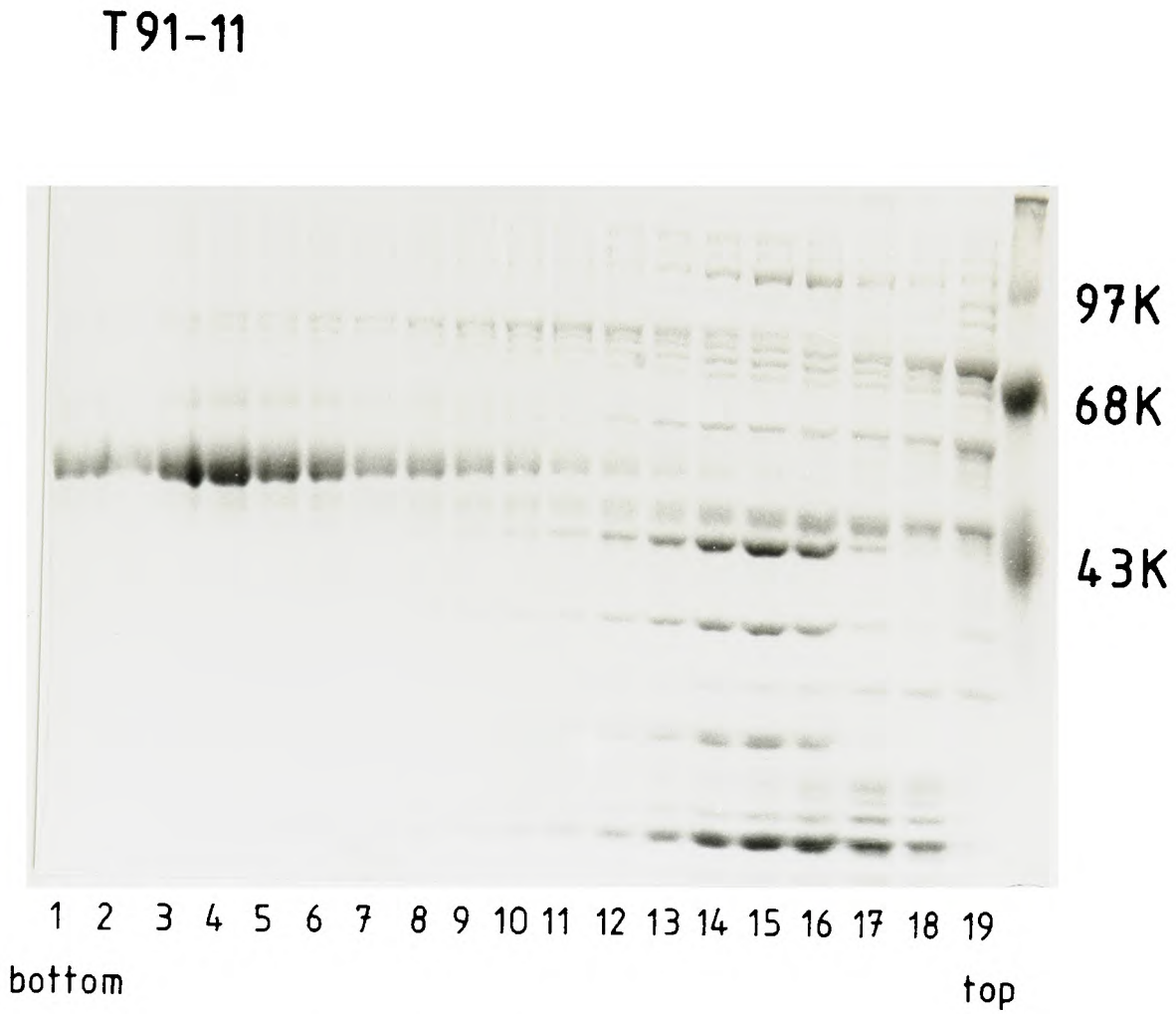
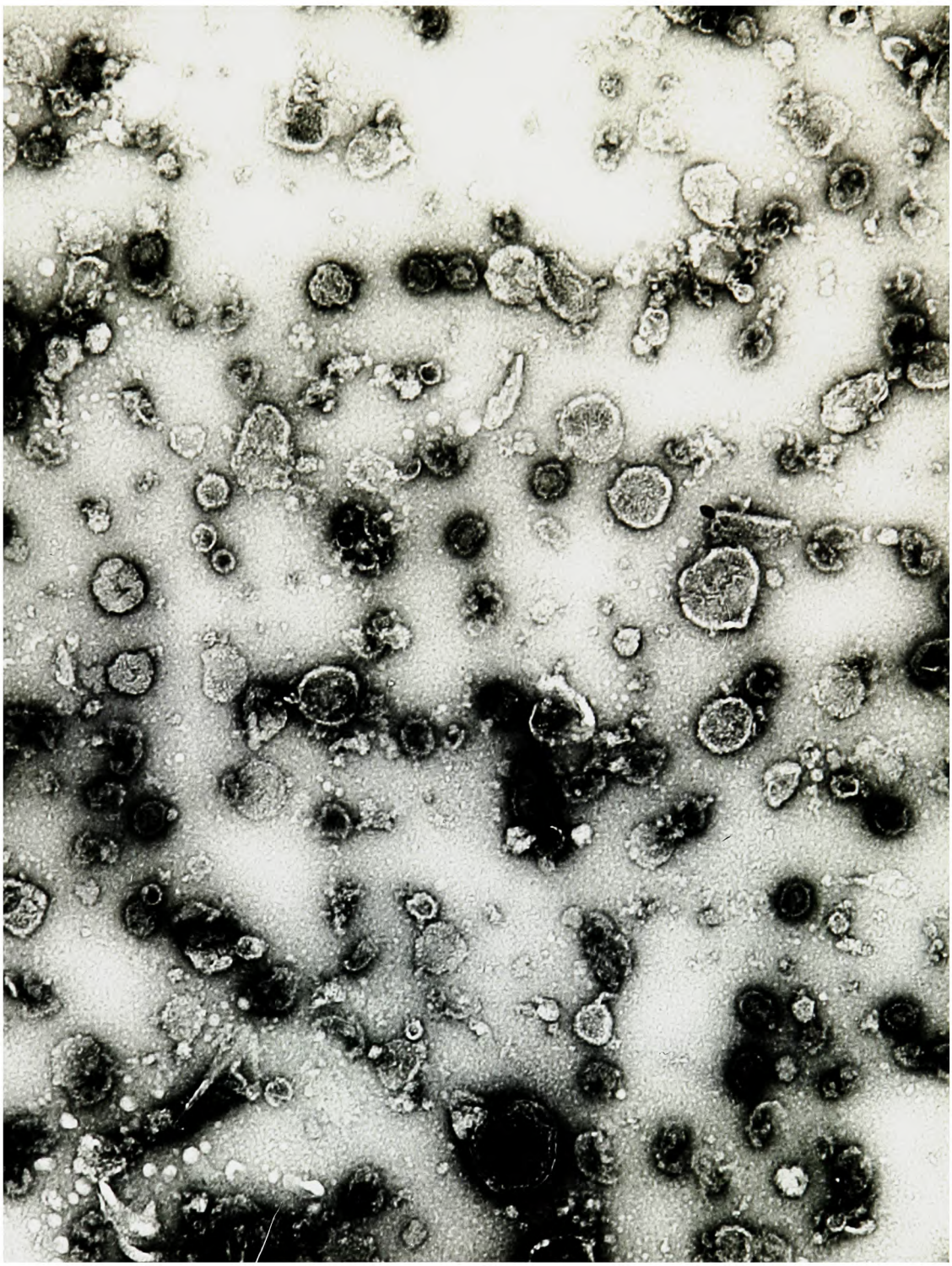


Figure 3.8. Coomassie Blue-stained gradient profile of a T91-11 extract. The preparation and treatment of the samples and the gel were described in legend to Figure 3.5.

Figure 3.9. Electron micrograph of uranyl acetate-stained Ty-VLPs from the pelleted fractions of a SW27 sucrose gradient fractionated T91-11 extract (X 30,160).



responsible for the proteolytic cleavages of p1 and p3 is a Ty encoded acid protease that is situated at the 5' end of TYB (Figure 1.4). A small in-frame deletion of this consensus sequence from the 1Ø fragment resulted in p1 and p3 production alone, with the Ty-VLPs having the same morphology as those in Figure 3.9, but no reverse transcriptase activity.

A combination of the results obtained in Sections 3.2, 3.3 and by Adams et al. (1987) has led to the hypothetical model for the life cycle of a Ty element that is outlined in Figure 3.1Ø. The integrated Ty element produces full length 5.7 kb RNA that is translated both simply and by frameshifting to produce the Ty proteins p1 (5Ø kD) and p3 (19Ø kD). These then aggregate, along with some 5.7 kb Ty RNA, and probably some $tRNA^{met}$, to form Ty-VLPs that are unstable and have the T91-11 particle morphology. The Ty encoded acid protease (part of p3) is then activated, and it cleaves p1 and p3 to yield the smaller proteins p2, p4, p5, p6 and reverse transcriptase. Consequently, the Ty-VLPs now have p2 as their major protein, are more stable and acquire the morphology of T91-1Ø particles. The simultaneity of the changes in particle morphology with proteolytic processing is reminscent of the maturation pathway of Murine Leukaemia Virus, whereby the initial particle complexes become infectious upon the proteolytic processing of the precursor proteins (Katoh et al., 1985; see Section 1.3.3.1). The VLP associated reverse transcriptase is now juxtapositioned to its Ty RNA template and $tRNA^{met}$ primer, so that a double stranded DNA copy is then produced. This may be in the form of a 2-LTR linear or circular cDNA which is then able to integrate into the genome, thus constituting a transposition

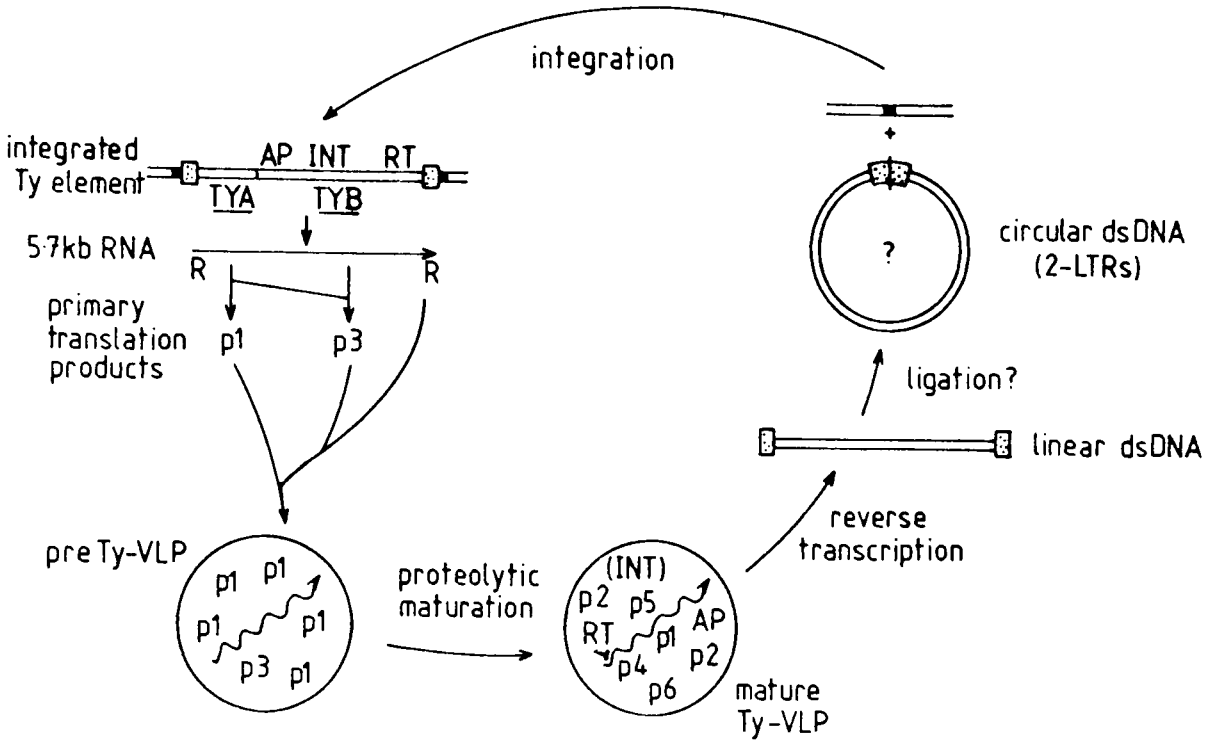


Figure 3.1Ø. The proposed life cycle of a Ty element. Double lines, DNA sequences (double stranded); solid boxes, target site DNA for transposition; stippled boxes, LTRs; AP, INT and RT, homologies to the consensus sequences of retroviral acid proteases, integrases and reverse transcriptases; single lines, RNA; R, terminal redundancy of 5.7 kb Ty RNA; p, protein; ↓, tRNA primer for reverse transcription; ?, proposed species and processes.

event. It should be stressed that although Ty transposition does occur via RNA (Boeke et al., 1985) no DNA transposition intermediates or Ty-VLP associated integrase activity have been identified to date.

The proposed Ty life cycle (outlined in Figure 3.10) is therefore very similar to that of a retrovirus (Figure 1.5). The major difference being that retroviral infections are intergenomic (from one cell to another) whilst Ty transposition is intragenomic (from one part of a given genome to another); the deciding factor for the choice of target genomes may be the presence or absence of an env-protein in the viral particles. It remains to be seen whether Ty-VLPs are essential for Ty transposition or whether they represent an inherited packaging pathway that isolates the potentially chaotic reverse transcriptase activity from the rest of the cytoplasm.

It has been suggested by Fink (1987) that the homologous recombination of intron containing yeast genes, with the cDNA copies of their spliced mRNAs, has led to the eradication of the intervening sequences from most yeast genes. It is proposed that the Ty encoded reverse transcriptase was responsible for producing the cDNA copies required for recombination. This simplistic view seems a little far-fetched since prokaryotes (with the exception of archaebacteria), that have no intron containing genes at all, have no access to reverse transcriptase. The lack of prokaryotic introns has been attributed to the phenotypic selection for the elimination of excess DNA that would, during DNA replication, utilise a large proportion of the

organism's total energy expenditure. This process is known as streamlining and is likely to have played a major part in the removal of introns from most yeast genes.

CHAPTER 4. RNA HOMOLOGOUS TO THE TYA REGION OF TY1-15 IS PACKAGED INTO TY VIRUS-LIKE PARTICLES : THE DEVELOPMENT OF AN ASSAY SYSTEM

4.1 Introduction: Retroviral RNA Packaging and its Exploitation

In Chapter 3 it was shown that Ty encoded proteins have the ability to form virus-like particles (Ty-VLPs). Like retroviral particles, Ty-VLPs contain RNA homologous to the element and have an associated reverse transcriptase. The presence of specific RNA packaging signals, that are essential for the inter-cellular transmission of retroviral sequences, within the genomic RNA of retroviruses (Section 1.3.3.1) suggests that there may be similar sequences involved in the packaging of the 5.7 kb Ty RNA into Ty-VLPs. Furthermore, it is likely that such sequences would be similarily essential for the transposition of Ty elements.

The factors that determine whether a retroviral RNA species is packaged into virions are *cis*-acting sequences that reside within the RNA itself. Some of these sequences are spliced out of the genomic RNA during the post-transcriptional processing that produces the sub-genomic RNAs, thus the genomic RNA is packaged whilst the *env*-mRNA is not. The assembly of retroviral proteins into particles does not require the packaging of RNA, therefore the analyses of retroviral mutants that produce low titres of infectious virus upon infection (or transfection with naked DNA, *in vitro*) but normal levels of virus particles have defined the signals responsible for the packaging of RNA into virus particles. The low infectivity of such mutants is a direct consequence of the reduced levels of retroviral RNA within the virions, resulting in a relatively small number of genomes being available for transmission by retroviral infection.

The packaging signals of Moloney Murine Leukaemia Virus

(MoMLV) and Spleen Necrosis Virus (SNV) have been localised to the region of the leader sequence that is 5' to the gag gene but 3' to the putative env-mRNA 5' splice site (Mann et al., 1983; Schwartzberg et al., 1983; Mann and Baltimore, 1985; Watanabe and Temin, 1982). The deletion of a 350 bp sequence, known as ψ, from the MoMLV leader sequence results in a 3,000 fold drop in the titre of infectious virus (Mann et al., 1983; Mann and Baltimore, 1985). Other regions of the genome are also involved in RNA packaging since the deletion of ψ does not totally prevent infectivity. Indeed, Bender et al. (1987) have recently shown that regions within the gag gene of MoMLV are required for efficient packaging, whereas Mann and Baltimore (1985) have demonstrated the flexibility of ψ, in that its relocation elsewhere within the MoMLV genome results in the partial retention of its ability to direct packaging.

The packaging requirements for avian retroviral RNAs are somewhat different to those of MoMLV and SNV because the putative env-mRNA 5' splice site is within the p19 region of the gag gene such that the 5' leader sequences are common to both the genomic and env-mRNA species. Regions downstream of the 5' splice site but still within the p19 region have been shown to be important for the RNA packaging of RSV (Pugatsch and Stacey, 1983; Darlix, 1986) and Avian Erythroblastosis Virus (Sealy et al., 1983), although the precise location of the nucleotides involved is rather controversial (Norton and Coffin, 1985). Sequences within the 5' non-coding region, however, are also involved; the naturally occuring mutants of RSV and Avian Sarcoma Virus (ASV),

TK15 and SE21Q1B respectively, both have large deletions in this region and are deficient in packaging (Koyama et al., 1984; Nishizawa et al., 1985; Shank and Linial, 1980). In a subsequent study, a region that is conserved within the leader sequences of various avian retroviruses has been shown to be required, but not to be sufficient, for the packaging of ASV RNA; its deletion restricts genomic RNA packaging but its presence does not direct the packaging of the env-mRNA (Katz et al., 1987). Sequences within the LTRs of RSV (Darlix, 1986) and the direct repeats that flank the src gene of RSV (Sorge et al., 1983) have also been implicated in packaging. These data indicate that different regions of a given RNA molecule from either mammalian or avian retroviruses are recognised, possibly as a specific complex RNA structure, as the "packaging signal".

There are therefore two predominant fates for the genomic RNA of retroviruses. They can be either translated (with or without splicing) to produce viral proteins or assembled, along with the viral proteins, into viral particles. These two events need not be mutually exclusive, for instance, an mRNA molecule might act as a template for protein synthesis and then assemble, along with the newly formed proteins, into particles. Thus, these two events could be regarded as being coupled since the particle forming proteins and the ribosomes would be competing for the binding of the RNA. However, once an RNA molecule has been packaged, it ceases to be a template for protein synthesis; these two processes need to be carefully regulated if a retroviral infection is to be efficient.

Packaging mutants of MoMLV, that have had Ψ deleted, have

been used to construct stably transformed cell lines that produce high levels of viral particles containing no retroviral RNA (Mann et al., 1983). If such a cell line is transfected with a plasmid that produces an RNA containing the ψ sequence, but lacking other retroviral sequences, that RNA will be incorporated into the viral particles. These virions can be harvested and used to infect a second cell line at high efficiency. The packaged RNA is therefore introduced into cells where it can be reverse transcribed into dsDNA, thus allowing genomic integration and the subsequent establishment of a stably transformed cell line. As a result, heterologous genes can be effectively introduced into cells by retroviral infection (Figure 4.1). Retroviral vector systems based on the above theory have been developed (Cepko et al., 1984; Miller et al., 1986) and constitute a possible route for gene therapy (Belmont et al., 1986; Hock and Miller, 1986; Sorge et al., 1987) and a useful experimental tool (Jat et al., 1986).

4.2 Experimental Approach to the Analysis of Ty RNA Packaging Signals

It has been shown that the 5.7 kb Ty RNA (genomic RNA) is efficiently packaged into Ty-VLPs (Figure 3.3). There are two alternative experimental approaches that could be adopted to determine whether or not the packaging of Ty RNA into Ty-VLPs is a pre-requisite for Ty element transposition. Firstly, the transposition of mutated Ty elements, in the presence of an

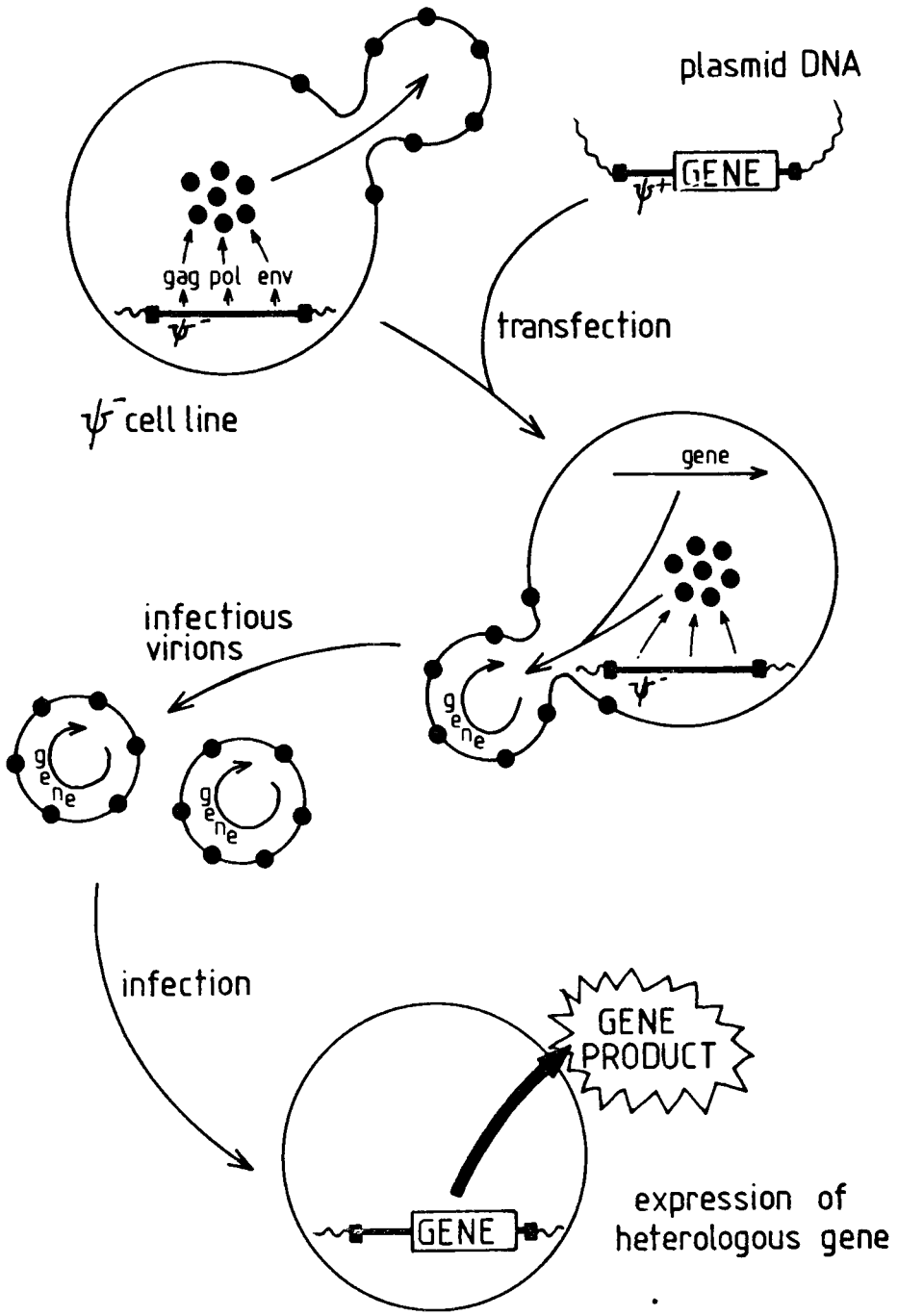


Figure 4.1. Schematic representation depicting the theory of retroviral vector systems. Large circles, cultured animal cells; small circles, retroviral particles; solid circles, retroviral proteins (gag, pol and env gene products); thick lines, unique retroviral DNA sequences that are chromosomally integrated or present in plasmids; solid boxes, retroviral LTRs; thin arrows with "gene" written nearby, RNA species. For a fuller explanation, refer to the text.

expression-competent "helper" element, could be measured. The "helper" element would ensure that an abundance of authentic Ty proteins, and therefore Ty-VLPs, would be present; these can be described as the trans-factors that are required for packaging. The drawback of such an approach is that the inhibition of transposition could be due to a number of factors other than defective packaging itself. For example, a failure of either the RNA to reverse transcribe or the cDNA to integrate would both prevent transposition. Secondly, the direct isolation of Ty-VLPs from cells containing both manipulated and expression-competent Ty elements, followed by the analysis of their RNA content would define the cis-acting regions important for RNA packaging. The effect of the disruption of such sequences on transposition could then be studied.

The second approach was adopted in an attempt to delineate the packaging signals of the class I element Ty1-15. An assay system was developed that consisted of two components. These were both high copy number plasmids with different selectable markers that could be cotransformed into yeast under double selection; they are referred to as the "helper" and "assay" plasmids. The same "helper" plasmid was used in all of the assays and is called pMA132-10. It is the same as pMA91-10 (Figure 1.8) except that the LEU2 selectable marker has been replaced by the TRP1 gene. This plasmid contains the whole transcriptional unit of Ty1-15 under the control of the powerful PGK promoter; high levels of the 5.7 kb Ty RNA, Ty encoded proteins and therefore Ty-VLPs are produced. The second plasmid, or "assay" plasmid, was different in each experiment; its

selectable marker was the LEU2 gene so that its cotransformation into yeast with the "helper" plasmid was assured. The "assay" plasmids were all designed so that the RNA species to be tested was "tagged" with sequences homologous to the human α2-interferon gene. The "tag" enables the "assay" RNA to be distinguished from all the other RNA species of yeast because the yeast genome only has very faint homology to interferon (data not shown). It was known that some of the "assay" RNAs would produce particle-forming proteins themselves (for example, RNA from the 11 fragment, Figure 3.8). The "helper" plasmid was, however, present in all the assays to ensure that plenty of mature Ty-VLPs were also available for RNA packaging.

It was shown in Chapter 3 that the 5.7 kb Ty RNA was packaged into Ty-VLPs, such that a clear peak of Ty RNA, corresponding to its association with particles, was observed when the RNA samples from a fractionated Ty-VLP gradient were analysed by Northern blotting (Figure 3.3). The packaging capabilities of various IFN tagged RNAs ("assay" RNAs) were therefore assessed in a similar manner. The presence of some or all of the cis-requirements for packaging, within a given "assay" RNA, was demonstrated by that RNA having a peak that coincided with the peak of the 5.7 kb Ty RNA ("helper" RNA).

4.3 RNA from the 5' end of Ty1-15 is Packaged into Ty-VLPs

As a preliminary packaging experiment, four different interferon (IFN) tagged RNAs were tested for their ability to be

packaged into Ty-VLPs by the approach that was outlined in Section 4.2. The regions of these four "assay" plasmids that produced the hybrid RNAs are shown in Figure 4.2, along with the "helper" plasmid; the selectable marker for each plasmid is also indicated. The plasmids pMA91-11-8, pKT346-8 and pKT219-8 all contain the TYA region of Ty1-15 in the form of the 11 fragment (Figure 1.8) whilst pMA943 has no Ty homologous sequences. The initial experiments were focused on these plasmids because of the expectation that RNA corresponding to the 5' end of a Ty element, like that from retroviruses, ought to contain some packaging signals. The IFN tagging of all four RNAs was achieved by using the human interferon α2-cDNA fragments, known as IFN-1 and IFN-8, in the plasmid constructions (Figure 4.3; Mellor et al., 1985d). The IFN-1 fragment contains a fortuitous yeast terminator which has been deleted in the IFN-8 fragment. The plasmids pMA91-11-8 and pKT346-8 both use the powerful PGK promoter for driving expression whilst pKT219-8 uses the authentic Ty expression signals. The termination of transcription is mediated by the PGK terminator in pMA91-11-8 and by the 3' LTR of Ty1-15 in pKT346-8 and pKT219-8. The plasmid pMA943 contains the PGK promoter, the PGK coding region and the IFN-1 fragment at the +4 position with respect to the termination codon of PGK; therefore the IFN sequences are responsible for the termination of transcription. These four "assay" plasmids were cotransformed into the yeast strain MD40-4c with pMA132-10 using double selection. For the convenience of nomenclature, all yeast strains that are cotransformants of a given plasmid, pX, and pMA132-10 will be referred to as PTX (Packaging Transformant X).

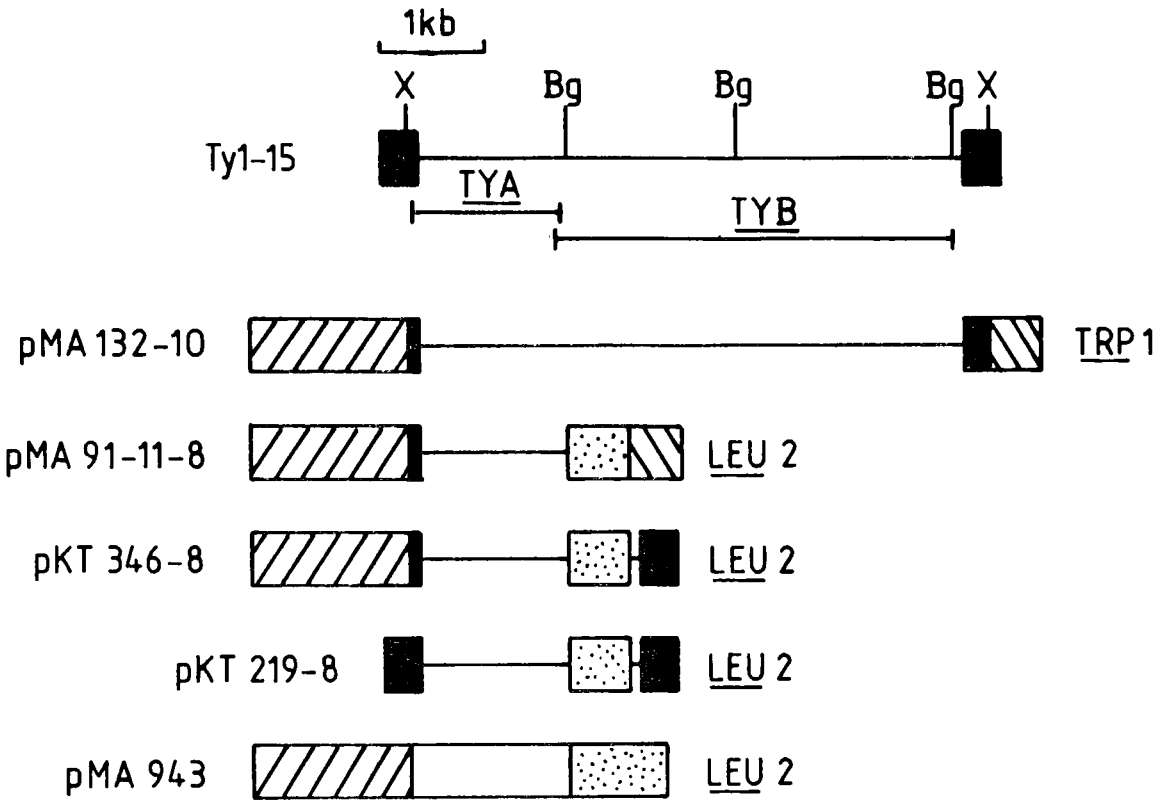


Figure 4.2. Diagram showing the genome organisation of Ty1-15 and the regions of the plasmids pMA132-10, pMA91-11-8, pKT346-8, pKT219-8 and pMA943 that produce the Ty, Ty-IFN or PGK-IFN RNAs. The selectable marker (LEU2 or TRP1) is indicated for each plasmid. Solid boxes, LTR derived sequences; thin lines, unique sequences of Ty1-15; hatched boxes, 5' and 3' regions from the yeast PGK gene; stippled boxes, IFN sequences; open box, PGK coding region; X, XhoI; Bg, BglII.

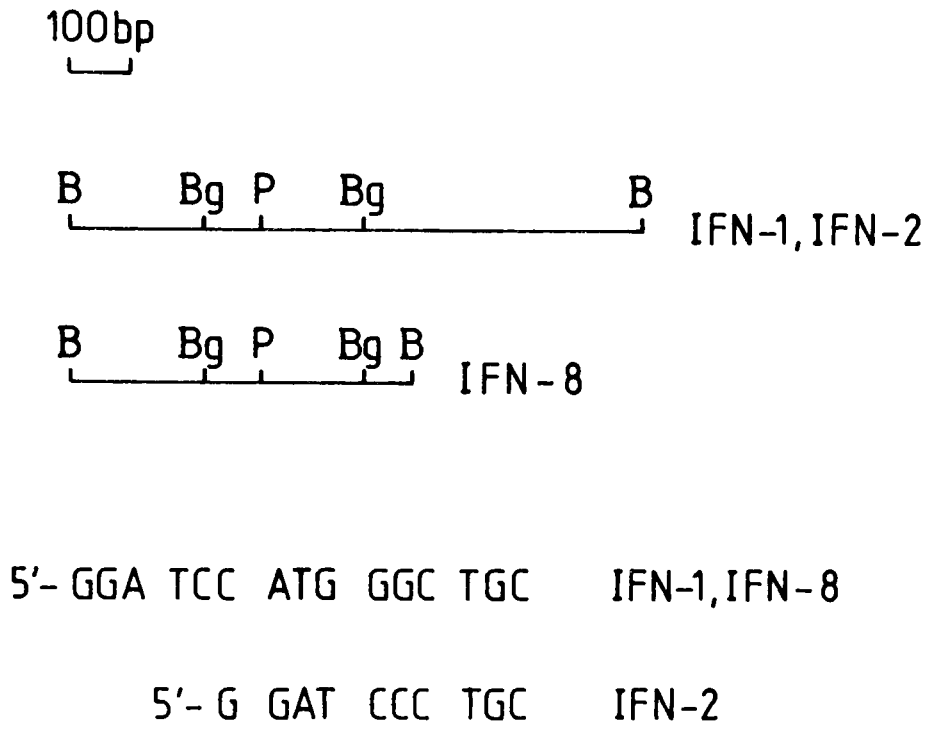


Figure 4.3. The restriction maps and 5' termini (containing the engineered BamHI sites) of the IFN-1, IFN-2 and IFN-8 fragments (Mellor et al., 1985a; Mellor et al., 1985d). B, BamHI; Bg, BglII; P, PvuII.

Mini VLP extracts were prepared from selective cultures of PT91-11-8, PT346-8, PT219-8 and PT943 and fractionated in SW40 15-45% (w/v) sucrose gradients. The RNA from each fraction was isolated and analysed by Northern blotting (as described in Section 2.18.2). The filters were all hybridised with an IFN specific probe to determine the position, within the gradient, of the "assay" RNAs. Following autoradiography, the filters were reprobed with the middle BglII fragment (1.65 kb) from Ty1-15 (Figure 4.2): this probe was specific for the 5.7 kb Ty RNA ("helper" RNA). The gradient profiles of all four IFN-probed filters and the Ty-reprobed PT91-11-8 filter are aligned in Figure 4.4. Only one 5.7 kb Ty RNA profile is shown because all four such profiles were, in fact, the same. There is a clear peak of the 5.7 kb Ty RNA between fractions 4 and 9, indicating the presence of Ty-VLPs in this part of the gradient; this correlates well with the data in Figure 3.3. There are similar peaks of IFN homologous RNA ("assay" RNA) in the fractionated PT91-11-8, PT346-8 and pT219-9 gradients, demonstrating that these RNAs are also packaged into Ty-VLPs. The virtual absence of the PT943 "assay" RNA from the Ty-VLP containing fractions demonstrates that this RNA is not packaged into Ty-VLPs. It therefore appears that the packaging of Ty RNA is a specific event and that at least some of the necessary information resides within the first 1,465 nucleotides of the 5.7 kb Ty RNA (in other words, the region that is homologous to the 11 fragment).

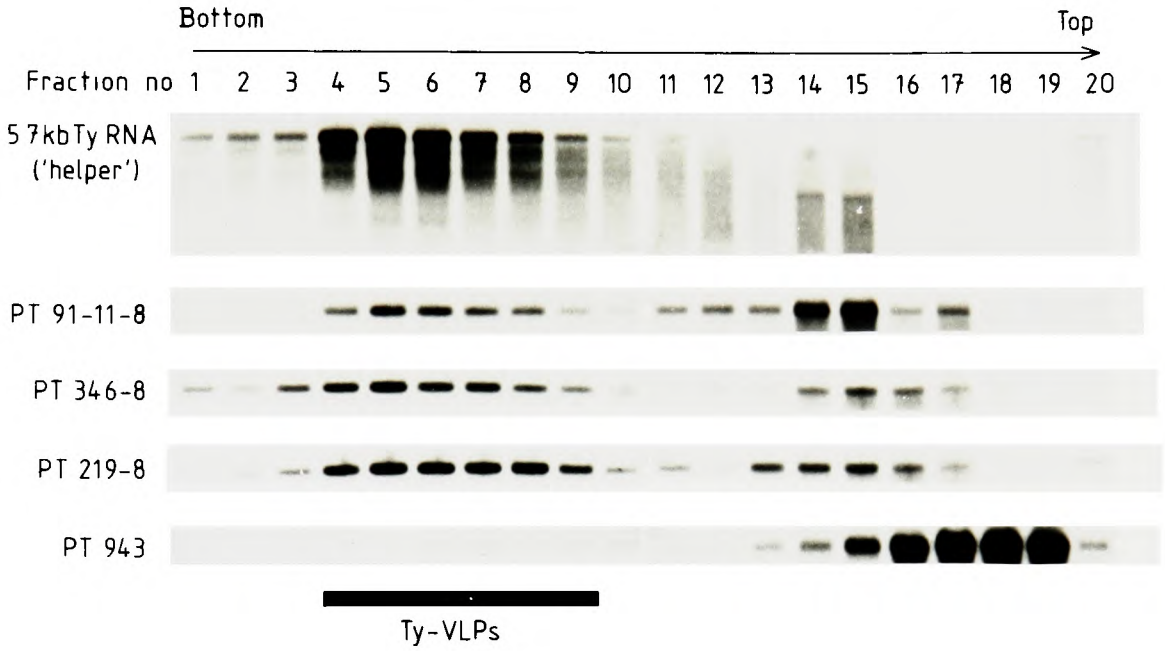


Figure 4.4. Northern blot analyses showing the RNA profiles of SW4Ø sucrose gradient fractionated extracts of PT-91-11-8, PT346-8, PT219-8 and PT943. The fractions previously shown to contain Ty-VLPs are indicated. The hybridisation probes were either the IFN-1 fragment ("assay" RNA specific) or the 1.65 kb BglII fragment corresponding to the 5' region of TYB (refer to Figure 4.2) ("helper" RNA specific).

4.4 The Evaluation of Packaging Directed by the 11 Fragment

The data in Figure 4.4 showed that RNA corresponding to the 11 fragment of Ty1-15 contains sequences that direct the specific packaging of itself into Ty-VLPs. The obvious drawback with this observation is that it is only qualitative, the only question it answers is whether a given "assay" RNA is packaged or not. The analyses of retroviral packaging signals by the measurement of the titres of infectious virus were quantitative, and it was shown that certain mutations could alter the infectivity of MoMLV by several orders of magnitude (Section 4.1). Consequently, it was considered important to quantitatively assay the RNA content of Ty-VLPs.

The most direct approach for quantitating the packaging of the "assay" RNAs is to determine what proportions of those RNAs are packaged, in the context of the total RNA content of those cells. This was done by extracting the total cellular RNA from part of the same yeast culture from which the VLPs for sucrose gradient fractionation were prepared, and running a small amount of it (1-5 μg) adjacent to the pooled RNAs from the previously identified Ty-VLP containing fractions (numbers 4 to 9, Figure 4.4) in a Northern gel, followed by transfer to nitrocellulose. The resulting filter was simultaneously probed with both the IFN and Ty specific probes so that both the "assay" and "helper" RNAs were detected. After autoradiography, the bands were cut out of the filter and counted by liquid scintillation (Section 2.18.7).

The physical nature of the 5.7 kb Ty RNA is invariant in all of the assays. The majority of it (~90%) is always transcribed

from the same plasmid, pMA132-10, whilst the rest corresponds to the low levels that are produced by the 30-35 endogenous Ty elements. It can therefore be assumed to be packaged into Ty-VLPs with constant efficiency in all of these experiments. Consequently, the ratio of the "helper" RNA in the VLP (V) track (V_H) to the "helper" RNA in the cellular (C) track (C_H), acts as a loading control for the quantitation of the variable packaging of the "assay" RNA. The actual values obtained for the V to C ratios (V_H/C_H) of the "helper" RNA will obviously fluctuate in accordance with variable loading in both tracks; however, as the in vivo packaging efficiency of this RNA has been assumed to be constant, this value serves to standardise the V to C ratio (V_A/C_A) obtained for the "assay" RNA when divided into it. The results obtained from different experiments can be quantitated, and therefore compared.

The calculation that is outlined above can be represented by the following equation and the value that it generates is defined as the Packaging Coefficient (C^{pac}); the greater this value is, the more efficiently the "assay" RNA is packaged into Ty-VLPs. Since the C^{pac} value is only a measure of the proportion of the "assay RNA that is packaged, its magnitude is not affected by the cellular levels of the "assay" RNA.

$$C^{pac} = \frac{V_A \cdot C_H}{C_A \cdot V_H}$$

The C^{pac} values of the "assay" RNAs from PT91-11-8, PT346-8, PT219-8 and PT943 were therefore calculated by pooling the RNAs

from fractions 4 to 9 (Figure 4.4) and analysing them by Northern blotting along with the total cellular RNAs from the same cultures. The results obtained are shown in Figure 4.5, as are the C^{pac} values obtained. These four observations were repeated in entirety (data not shown) and the C^{pac} values were shown to be reproducible (Figure 4.5, 2^{nd} experiment).

There are two main conclusions that can be drawn from these results. Firstly, RNA that is homologous to the 11 fragment contains signals that direct its specific packaging into Ty-VLPs: RNA that contains no Ty homologous sequences is not packaged. The C^{pac} of Ø.1 for the non-packaged "assay" RNA of PT943 reflects the slight contamination of the whole sucrose gradient with this RNA, a consequence of minor association to large pieces of cell debris. Secondly, this data shows that the presence of RNA sequences that are homologous to the 3' LTR, rather than the PGK terminator, up to the point of transcription termination, contribute towards more efficient packaging (demonstrated by PT219-8 compared to PT91-11-8). It is possible, however, that the different levels of p1 produced from the 11 fragment "assay" RNAs could influence the outcome of the C^{pac} calculation. p1 has previously been shown to form particles (Figures 3.8 and 3.9), therefore the number of VLPs available for RNA packaging in the cytoplasms of the different strains will be different. The "assay" RNAs of PT219-8 (Ty promoter) and PT346-8 (PGK promoter) have similar C^{pac} values, but different p1 expression efficiencies. This argues against the generation of artefactual results by the variable levels of p1. There will be further data

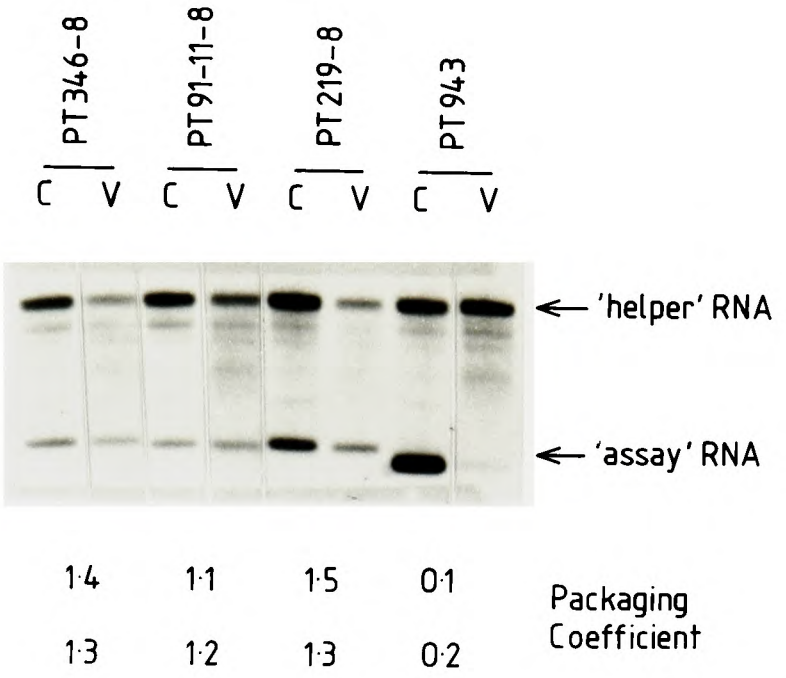


Figure 4.5. The Packaging Coefficients of the "assay" RNAs from PT346-8, PT91-11-8, PT219-8 and PT943. The autoradiograph corresponds to the Northern filter that was used for the determination of the C^{pac} values of the "assay" RNAs in Figure 4.4. The values obtained in the repeat experiment are also shown. The tracks labelled 'C' contain total cellular RNA, whilst those labelled 'V' contain the RNA from the Ty-VLP containing fractions. The hybridisation probes were described in the legend to Figure 4.4.

to support this conclusion in Section 4.6.

A feature of the results presented in Figure 4.5 is that some of the C^{pac} values are in excess of 1.Ø, indicating that a greater proportion of the "assay" RNA, as opposed to the "helper" RNA, appears to be packaged into the Ty-VLPs. It would be predicted that the 5.7 kb Ty RNA (genomic) would be the most efficiently packaged RNA species and that its C^{pac} value would be 1.Ø; however there are several possible explanations for this apparently paradoxical observation. Firstly, the "helper" RNA is not promoted and terminated by the authentic Ty 5' and 3' LTRs; it has already been shown that the presence of sequences homologous to the 3' LTR at the 3' terminus of the RNA contributes to increased packaging (Figure 4.5). The pMA132-1Ø produced RNA is therefore unlikely to be packaged with maximum efficiency. Secondly, the 5.7 kb Ty RNA is known to be particularly sensitive to RNase degradation (due to its large size) as are the VLP RNAs (due to a 3.5 - 4 hour gap between cell breakage and the addition of phenol); therefore the V_H component of the C^{pac} equation may be diminished by degradation and, since it is a denominator, its reduction results in increased C^{pac} values. Thirdly, it may be that smaller RNA species are packaged in preference to the larger 5.7 kb Ty RNA by virtue of each packaged molecule occupying less space. It is important to realise that the utility of the C^{pac} calculation is in giving a numerical comparison of the packaging efficiencies of various "assay" RNAs with each other, and not with the 5.7 kb Ty RNA itself.

4.5 The Construction of a 3' to 5' Deletion Series through the 11 Fragment

The packaging of RNA homologous to the 11 fragment of Ty1-15, into Ty-VLPs, has been demonstrated and quantitated (Sections 4.3 and 4.4). In an attempt to localise the regions that are responsible for this, a unidirectional 3' to 5' deletion series through the 11 fragment was constructed. The scheme for these plasmid constructions is summarised in Figure 4.6.

The starting plasmid was pKT227. This contains the origins of replication and the selectable markers for its propagation in both E. coli and yeast. It also contains both of the LTRs of Ty1-15 and the regions of epsilon that are upstream of the first BglII site and downstream from the third BglII site; it therefore contains the whole of TYA and the authentic Ty expression signals. The direction of Ty transcription in pKT227 is indicated in Figure 4.6.

The deletion series was constructed by digesting pKT227 at the unique BglII restriction site and treating the linearised DNA with Bal31 exonuclease (as described in Section 2.9.6). Aliquots of the reaction were removed at one minute intervals over a period of 8 minutes. The exonuclease digested DNAs were filled-in, blunt-end ligated in the presence of excess synthetic BamHI linkers (CCGGATCCGG; Pharmacia) and transformed into E. coli. The extent of the nuclease digestion and the insertion of linkers were anlaysed by preparing "miniprep" DNA from the transformants and digesting them with BamHI and HindIII (pKT227 has a unique HindIII site). Figure 4.7 shows a typical sample of minipreps

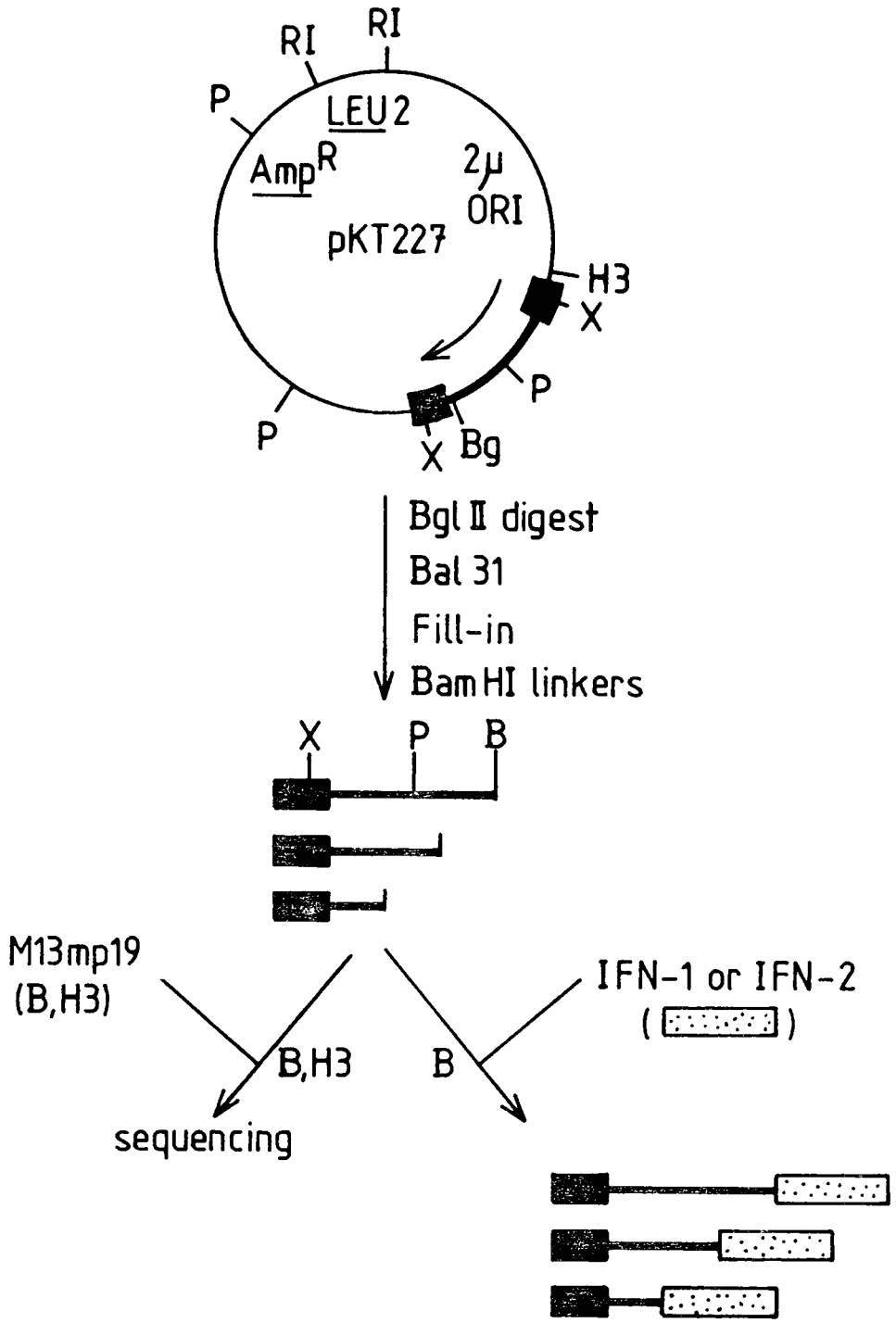


Figure 4.6. The construction of the 3' to 5' deletion series through the 11 fragment starting from pKT227. Thin lines, yeast/*E. coli* selection and replication sequences (similar to those of pMA91, Figure 1.8) and unspecified yeast sequences flanking Ty1-15; solid boxes, LTRs; thick lines, Ty1-15 unique sequences (mainly *TYA*); stippled boxes, IFN sequences; the direction of Ty transcription within pKT227 is indicated by the arrow; RI, EcoRI; P, PstI; X, XhoI; Bg, BglII; H3, HindIII; B, BamHI.

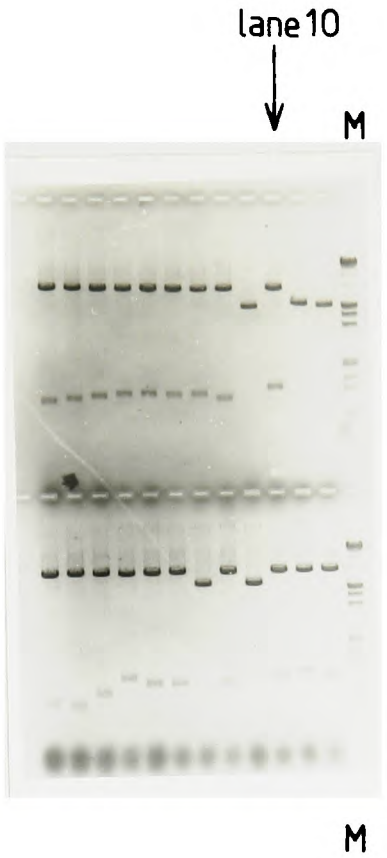


Figure 4.7. A 1% Tris-acetate agarose gel showing a typical sample of miniprep DNAs isolated from E. coli transformants of the pKT227 Bal31 deletion series, digested with BglII and HindIII. DNAs in lanes 1, 5, 13, 14, 15 and 20 were selected for endpoint determination. The size markers (M) are DNA digested with HindIII and EcoRI; the band sizes (in kb) are 21.7, 5.15, 5.00, 4.27, 3.48, 1.98, 1.59, 1.37, 0.94, 0.83 and 0.56.

from the 1,2,3 and 4 minute timepoints that were analysed in this way, lane 10 contains pKT227 DNA that was digested with BglII and HindIII. DNAs that had endpoints in the required regions were digested with BamHI and HindIII and ligated into BamHI,HindIII digested M13mp19 DNA for the determination of the deletion endpoints by DNA sequencing. The IFN "tags" were then inserted, as BamHI fragments, into the unique BamHI sites of the selected endpoint molecules so that the Ty and IFN reading frames were in phase. This was done in an attempt to guard against the premature termination of translation, a phenomenon that can result in the reduction of RNA stability (Losson and Lacroute, 1983). IFN fragments with two slightly different 5' termini (IFN-1 and IFN-2, Figure 4.3, Mellor et al., 1985a) were used. Hence, two-thirds of the deletion endpoints were able to have the Ty and IFN sequences fused together in-phase. In addition, the IFN-1 fragment was inserted into the BglII site of pKT227 to generate pKT227-1, whilst the IFN-2 fragment was inserted into the BglII site (at a converted PvuII site towards the 5' end of the TYA gene of Ty1-15) of pKT202 to generate pKT202-2. Figure 4.8 shows the final deletion series, the numbers of the plasmids and the IFN fragments that were used. Figure 4.9 shows the deletion series plasmids, with the exception of pKT202-2, digested with BglII and HindIII; the IFN fragments contain two BglII sites (Figure 4.3), consequently, the variable length of the Ty sequences can be visualised easily. All seven plasmids were cotransformed into MD40-4c with pMA132-10.

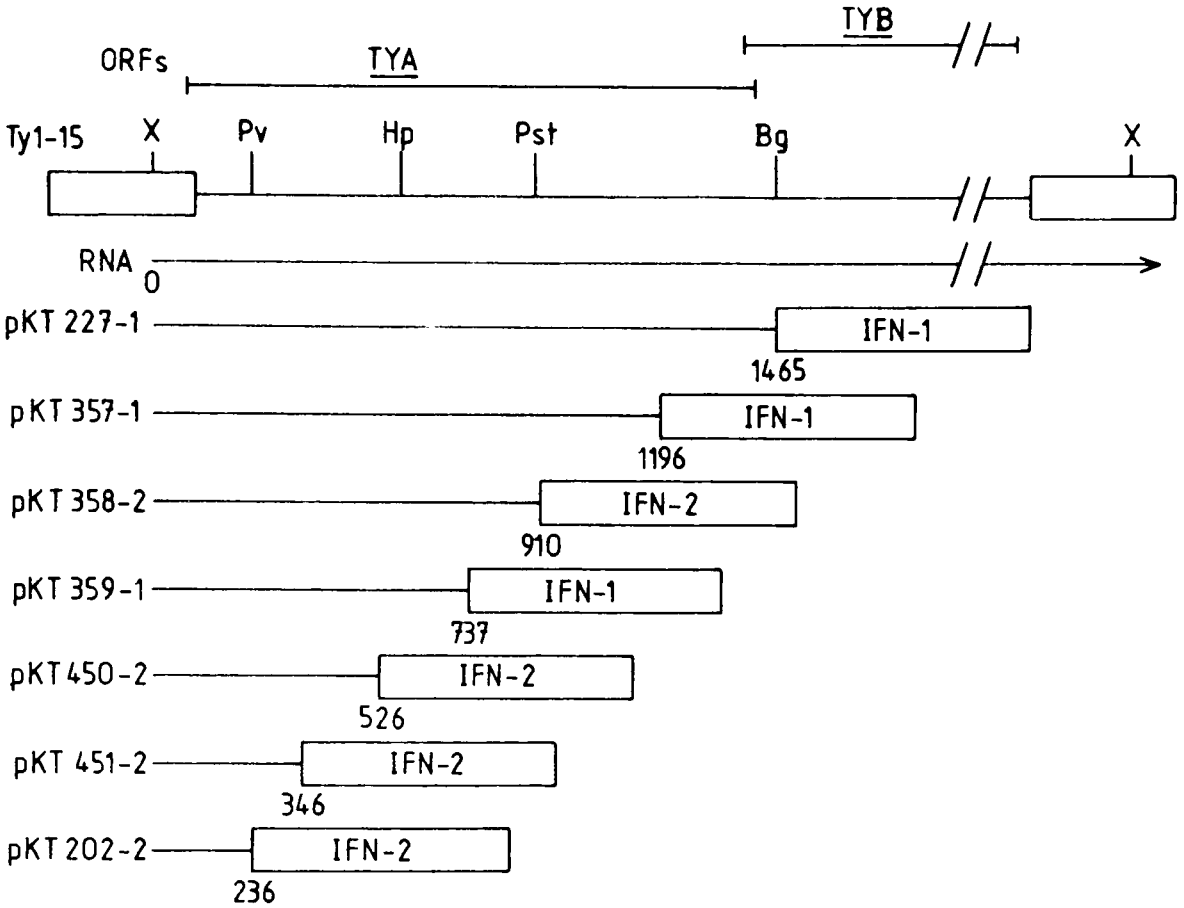


Figure 4.8. The restriction map of the TYA gene of Ty1-15 aligned with the regions of pKT227-1, pKT357-1, pKT358-2, pKT359-1, pKT450-2, pKT451-2 and pKT202-2 (the pKT227 deletion series) that produce the Ty-IFN hybrid RNAs. The numbers refer to the length of Ty homologous RNA from the start site of the 5.7 kb Ty RNA up to the first nucleotide of the BamHI linker. X, XhoI; Pv, PvuII; Hp, HpaI; Pst, PstI; Bg, BglII.

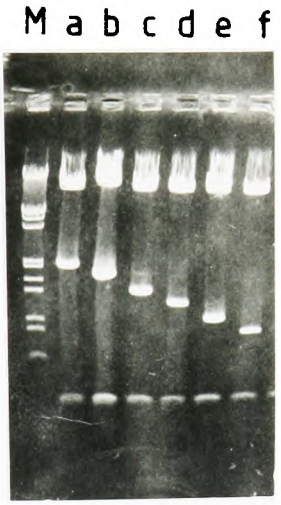


Figure 4.9. A 1% Tris-borate agarose minigel showing the digestion of pKT227-1 (lane a), pKT357-1 (lane b), pKT358-2 (lane c), pKT359-1 (lane d), pKT450-2 (lane e) and pKT451-2 (lane f) with BglII and HindIII. The markers (M) were described in the legend to Figure 4.7. The decreasing size of the middle DNA fragment represents the extent of the Bal31 digestion.

4.6 The Packaging Analysis of the 11 Fragment Deletion Series

The seven cotransformants, PT227-1, PT357-1, PT358-2, PT359-1, PT450-2, PT451-2 and PT202-2, that were made in Section 4.5, were assayed for their ability to package their IFN homologous RNAs into Ty-VLPs. The same approach that was used in Sections 4.3 and 4.4 was adopted. Figure 4.10 shows the profiles of the "assay" RNAs that were isolated from Ty-VLP extracts that had been fractionated on SW40 sucrose gradients. All the profiles have peaks of IFN homologous RNA in the fractions (approximately, numbers 4 to 9) that have been previously observed to contain Ty-VLPs. The presence of the 5.7 kb Ty RNA in these fractions was confirmed in all cases (data not shown). The isolated RNAs from the peak fractions of each of the seven gradients were pooled and analysed by Northern blotting in conjunction with total cellular RNAs that were extracted from the same cultures. The C^{pac} values of all seven "assay" RNAs were then calculated in the same way as before (Figure 4.11). All seven assays were repeated in entirety and the C^{pac} values that were obtained are also shown in Figure 4.11. These results are all presented graphically in Figure 4.12.

There are a number of direct observations that can be made with this data. The first one relates to the RNA levels in the 'C' tracks, rather than the packaging. If the level of 5.7 kb Ty RNA is assumed to be a loading control, the levels of the IFN "tagged" RNAs clearly diminish as the 11 fragment is sequentially deleted in the 3' to 5' direction. This implies that there are regions downstream from the Ty RNA start site that are important

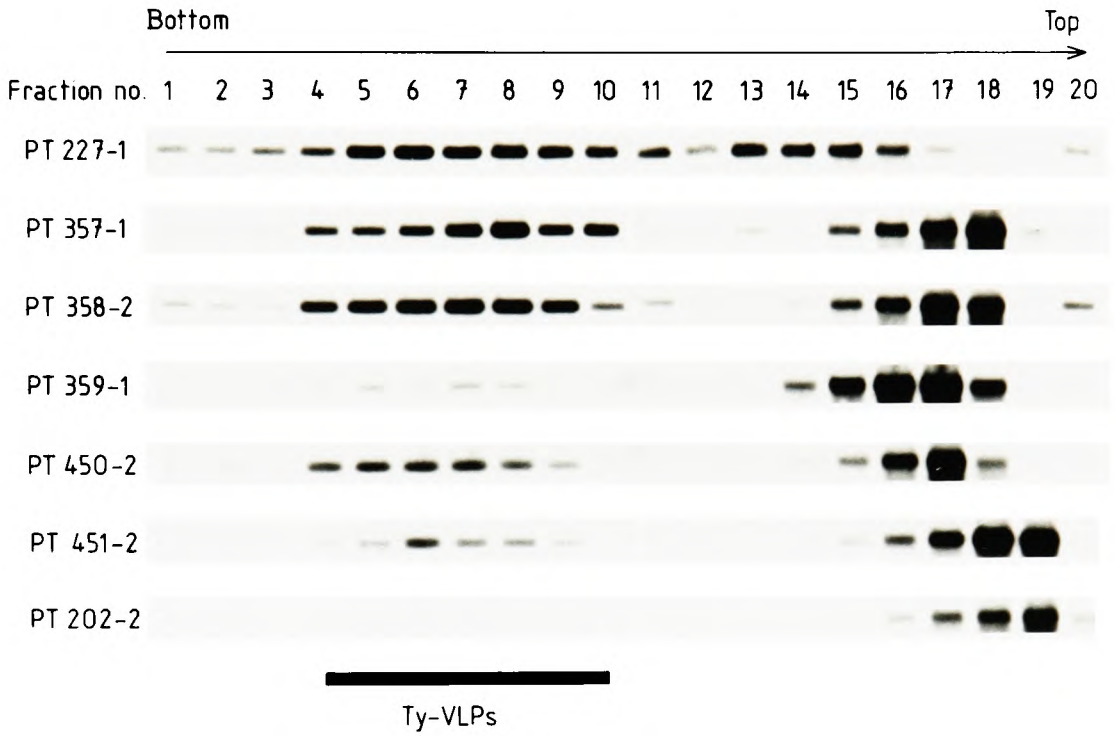


Figure 4.10. Northern blot analyses showing the "assay" RNA profiles of SW40 sucrose gradient fractionated extracts of PT227-1, PT357-1, PT358-2, PT359-1, PT450-2, PT451-2 and PT202-2. The fractions previously shown to contain Ty-VLPs are indicated. The hybridisation probes were described in the legend to Figure 4.4.

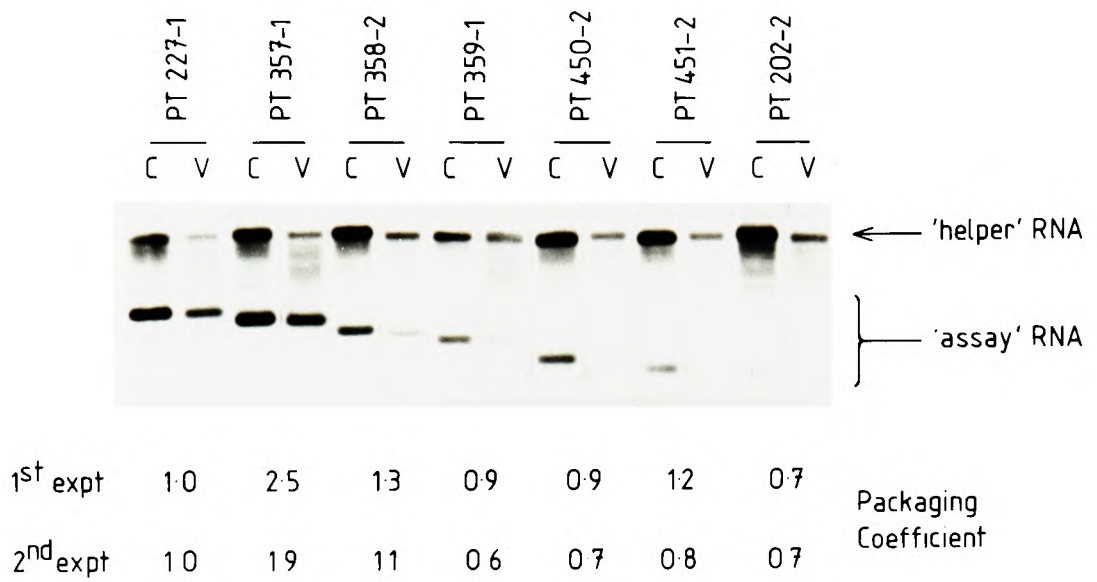


Figure 4.11. The Packaging Coefficients of the "assay" RNAs from PT227-1, PT357-1, PT358-2, PT359-1, PT450-2, PT451-2 and PT202-2. The autoradiograph corresponds to the Northern filter that was used for the determination of the C^{pac} values of the "assay" RNAs in Figure 4.10. The values obtained in the repeat experiment are also shown. The tracks labelled 'C' contain total cellular RNA, whilst those labelled 'V' contain the RNA from the Ty-VLP containing fractions. The hybridisation probes were described in the legend to Figure 4.4.

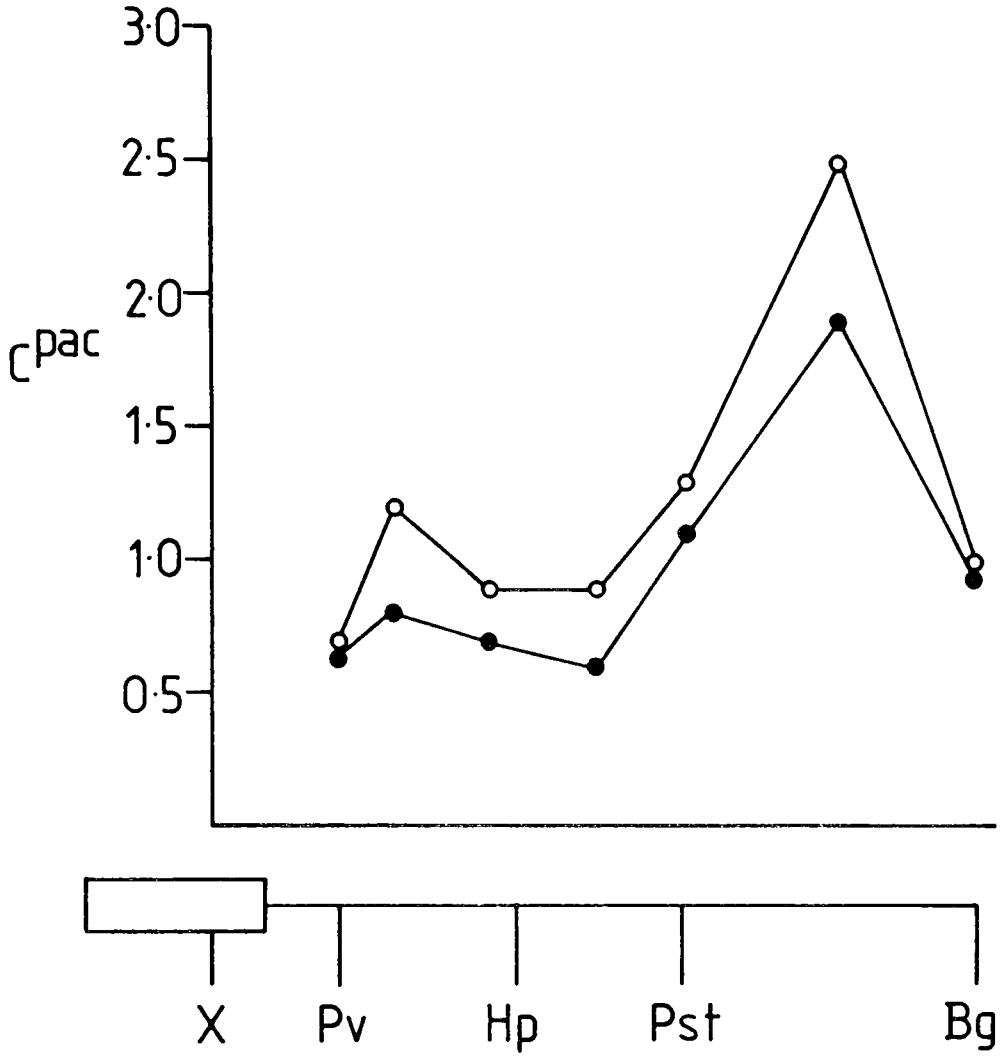


Figure 4.12. The alignment of the Packaging Coefficients for the "assay" RNAs from the 11 fragment deletions with the restriction map of the 11 fragment region of pKT227. Open circles, first experiment; solid circles, second experiment; X, XhoI; Pv, PvuII; Hp, HpaI; Pst, PstI; Bg, BglII.

for setting the levels of Ty transcription; this agrees with the data of Fulton et al. (Section 1.4.3), however, there are reservations about such a conclusion because there is no evidence that these RNAs have similar stabilities. For example, an RNA that is packaged efficiently is likely to accumulate because of the protection from RNases that is afforded by Ty-VLPs (refer to Section 3.4). Secondly, the C^{pac} value of the "assay" RNA from PT227-1 is both relatively small for an RNA that is homologous to the 11 fragment and similar to that of PT91-11-8 (1.Ø compared to 1.1). Both these RNAs have 3' termini that are not homologous with the 3' LTR: transcription termination occurs within the IFN homologous sequences in PT227-1 and within the PGK terminator in PT91-11-8. This confirms the earlier observation (Section 4.4) that the 3' LTR of Ty1-15 is important for efficient RNA packaging. It also supports the conclusion that the different levels of p1 produced from the "assay" RNAs do not influence C^{pac}; PT227-1 has levels of p1 resembling PT219-8, but has a considerably different C^{pac} (1.Ø compared to 1.5). Thirdly, the C^{pac} obtained with PT357-1 ("assay" RNA contains 1,196 nucleotides from the 5' end of the 5.7 kb Ty RNA) is reproducibly higher than that of PT227-1 (1,465 nucleotides of Ty RNA). This suggests that a sequence that is highly inhibitory to RNA packaging resides between the sequences that correspond to the BglII site of pKT227-1 and the pKT357 endpoint; alternatively, this may be due to some kind of biological artefact. For example, the PT357-1 "assay" RNA may adopt a complex secondary structure that culminates in its own efficient packaging or the disruption

of the "helper" RNA's packaging. The C^{pac} obtained with PT358-2 (910 nucleotides of Ty RNA) is also higher than that of PT227-1, similar arguments may therefore apply to this observation. Fourthly, the C^{pac} of PT359-1 (737 nucleotides of Ty RNA) is always less than that of PT227-1, PT357-1 and PT358-2, this implies that a positive packaging signal resides downstream from the first 737 nucleotides of the 5.7 kb Ty RNA. Fifthly, the C^{pac} of PT451-2 (346 nucleotides of Ty RNA) is greater than those of PT202-2 (236 nucleotides of Ty RNA) and PT450-2 (526 nucleotides of Ty RNA); this suggests that the PT451-2 "assay" RNA may also adopt a secondary structure that facilitates its association to VLPs, since the combination of a negative packaging signal between nucleotides 346 and 526, and a positive signal between nucleotides 236 and 346 seems improbable. Sixthly, the C^{pac} of PT202-2 is 0.7, this shows that the RNA sequences between the 5' terminus of the genomic Ty RNA and the sequence corresponding to the PvuII site of TYA (the first 236 nucleotides of the 5.7 kb Ty RNA) contain a positive packaging signal.

4.7 Discussion

This chapter has described an assay system that directly measures the packaging of an IFN "tagged" RNA into Ty-VLPs. The presence of an overexpressing "helper" Ty element in all of the assays ensured that suitable levels of the Ty proteins (the trans factors) were available for Ty-VLP formation. The cis-acting regions of a defective "assay" Ty element that, when transcribed, are responsible for RNA packaging can therefore be studied. The

results have been quantitated by using an equation that determines a numerical value termed the Packaging Coefficient (C^{pac}), thus the straightfoward comparison of different experiments was made possible.

The major observation of this chapter is that the packaging of a Ty RNA molecule is a specific event, and that some of the information responsible for this resides within the RNA sequences that are homologous to the 11 fragment (the region of Ty1-15 that contains TYA). The packaging analysis of a 3' to 5' deletion series through the 11 fragment was aimed at localising the sequences responsible for this. However, the results do not lend themselves to simple interpretation and, although reproducible, seem to be prone to biological artefacts that make it difficult to ascribe significance to the observations. Even so, it does appear that at least three different regions of the 5.7 kb Ty RNA, that is homologous to Ty1-15, contain positive signals for the packaging of RNA into Ty-VLPs. These are within the first 236 nucleotides (PT202-2 result), between nucleotides 737 and 1,465 (equivalent to the pKT357 endpoint and the BglII site of pKT227, respectively) and within the nucleotides homologous to the 3' LTR up to the point of transcription termination. The involvement of the RNA termini in RNA packaging might be analogous to the role of the RSV LTR homologous sequences in the packaging of genomic RSV RNA (Darlix, 1986). No other parallels between retroviral RNA and Ty RNA packaging, apart from its specificity, are discernible from these results.

It has been described how the variable levels of the VLP

forming protein, p1, produced by the 11 fragment homologous "assay" RNAs, could not have influenced the C^{pac} values obtained for those RNAs. This conclusion implies that any differential packaging efficiency between the "helper" and "assay" RNA species is not affected by the ratio of the p1-VLPs (T91-11 morphology) to the mature Ty-VLPs (T91-1Ø morphology) in the cell. In other words, the affinity of a given RNA to the particle forming proteins during particle (nucleoprotein) assembly is not influenced by the final morphology of that particle. In fact, this would be expected if the Ty life cycle (Figure 3.9) is correct. RNA is only packaged into VLPs with the T91-11 morphology during particle assembly, these are subsequently matured into VLPs with the T91-1Ø morphology with no accompanying change in their RNA content. However, it cannot be completely ruled out that excess p1 (produced from certain "assay" RNAs) still makes the comparisons of packaging with "assay" RNAs, that are not translated into a particle forming protein, somewhat uncontrolled. This situation might be further confused if some of the proteins that are produced from the truncated TYA genes, fused to IFN, also aggregate into particles; consequently, both authentic and engineered Ty proteins could be forming particles in the same cell, the results of which are unknown. Therefore, studies were undertaken to analyse RNA packaging signals using "assay" molecules that are incapable of producing any Ty protein (Chapter 5) and to delineate the TYA protein requirements for particle formation (Chapter 6).

CHAPTER 5. SEQUENCES THAT ARE CAPABLE OF DIRECTING RNA PACKAGING ARE DISPERSED THROUGHOUT TY1-15

5.1 Introduction

It was outlined in Section 4.7 that a modified system for assaying the packaging of Ty homologous RNA into Ty-VLPs was required. The approach of directly analysing the association of "tagged" RNAs with particles (Chapter 4) is a reasonable one, and since it is the properties of the "assay" RNAs that are being assessed, it seemed appropriate for the design of the "assay" plasmids to be altered. Any new system should satisfy a number of criteria. First, all the "assay" RNAs should be similar to each other, this will largely eliminate factors such as variable secondary structures. Second, if minimal variations in the sizes of the "assay" RNAs are going to be made by the insertion of variable regions of Ty, then the initial RNA size should be substantially large. Third, to make detection simple, the "assay" RNAs should be abundant. Fourth, the "assay" RNAs should have a suitable "tag", the use of IFN homologous sequences is as good as any. Fifth, the "assay" RNAs should be translated, and for consistency, the same protein should be produced in each case; this is thought to confer a degree of RNA stability (Section 4.5). Sixth, the protein that is produced from the "assay" RNAs should be non-particulate, so as not to interfere with the Ty-VLPs (supplied in <u>trans</u> by pMA132-10) during sucrose gradient fractionation. Seventh, the Ty homologous regions should be downstream from the stop codon of the translated region, so as to ensure their non-expression as part of a protein.

5.2 Plasmid Constructions

To satisfy the above criteria for a novel "assay" plasmid, pMA924 was constructed (Figure 5.1). This is basically the same as pMA943, which was described earlier (Figure 4.2). The only modification is the removal of the BamHI site at the 3' end of the IFN homologous sequences so that the remaining BamHI site that is sandwiched between the PGK and IFN sequences becomes unique. This manipulation was achieved by a partial BamHI digestion of pMA943, followed by a fill-in reaction and a blunt-end ligation. The unique BamHI site is therefore 4 bp downstream from the stop codon for PGK translation. The transcription terminator is within the IFN sequences (the IFN-1 fragment was used in the construction of pMA943). Thus, any fragment that is inserted into the BamHI site of pMA924 will be transcribed as a PGK-insert-IFN hybrid RNA, though only authentic PGK protein will be produced by its translation. It was shown earlier (Figures 4.4 and 4.5) that the PT943 "assay" RNA is not packaged into Ty-VLPs, it is ~2 kb in length and is a relatively abundant RNA species. pMA924 therefore satisfies all the criteria that were laid out in Section 5.1.

A variety of subgenomic Ty1-15 fragments were inserted into the BamHI site of pMA924, in the same transcriptional orientation as the PGK and IFN sequences, the details of which are illustrated in Figure 5.2 and defined in its legend. The 11a, 458 and 459 fragments were inserted directly by virtue of their compatible GATC sticky ends. The 452, 453, 454 and 457 fragments required an intermediate cloning step into either pSP46 or pSP56

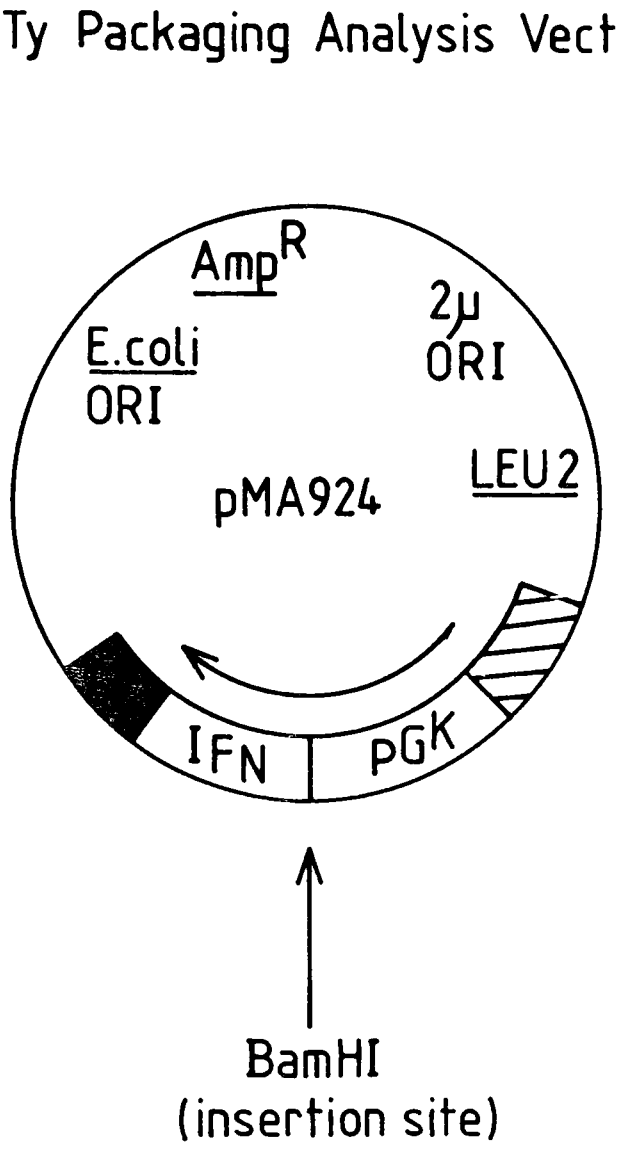


Figure 5.1. The structure of the Ty packaging analysis vector, pMA924. Thin line, yeast/E. coli selection and replication sequences (as indicated); hatched box, promoter sequences from the 5' region of the PGK gene; solid box, terminator sequences from the 3' region of the α2-IFN (cDNA) gene; open boxes, coding regions of the PGK gene and IFN sequences; the BamHI site for the insertion of sequences is indicated; the arrow indicates the direction of transcription of the PGK-insert-IFN hybrid RNA.

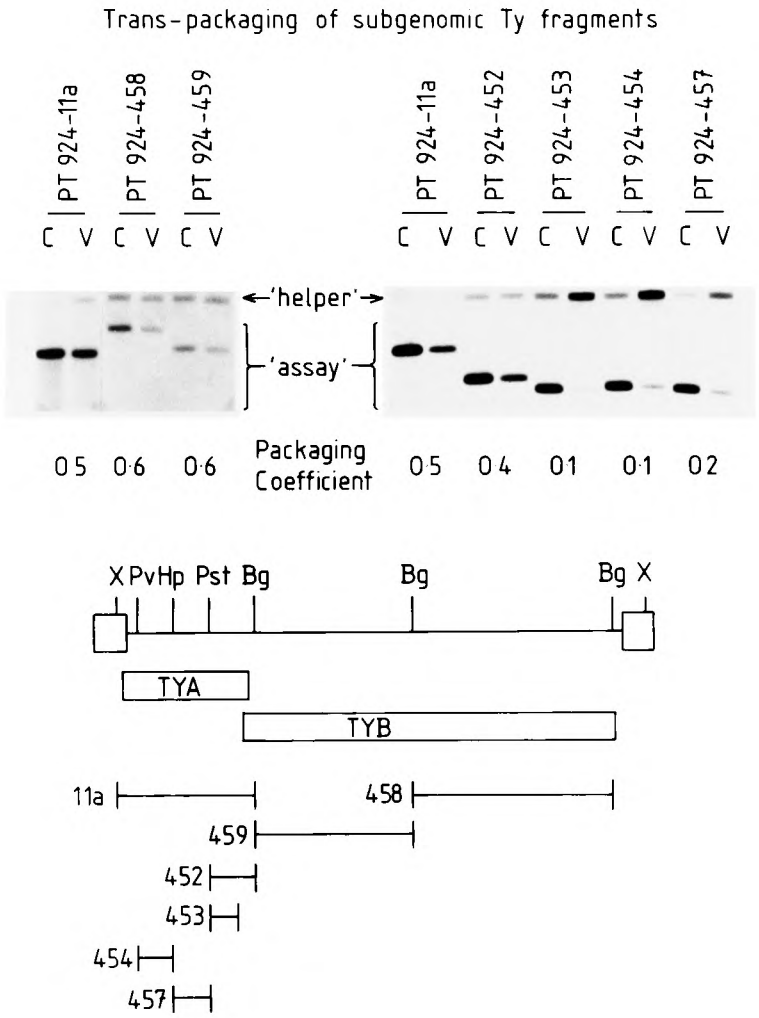
(Ogden et al., 1986). These two plasmids are derivatives of pSP64 and pSP65 respectively (Melton et al., 1984) where the HindIII sites have been converted to BglII sites. The fragments now had engineered BamHI and BglII termini so that their insertion and orientation determination in pMA924 was possible. All the derivatives of pMA924 were cotransformed into MD40-4c with pMA132-10 to produce the strains PT924-11a, PT924-452, PT924-453, PT924-454, PT924-457, PT924-458 and PT924-459.

5.3 Packaging Signals Reside Within RNAs Derived from Dispersed Regions of Ty1-15

The "assay" RNAs from the sucrose gradient fractionated extracts of the packaging transformants generated in Section 5.2 were analysed by Northern blotting (data not shown). The RNAs from the fractions corresponding to the Ty-VLP peaks were pooled and analysed on further Northern blots, adjacent to the total cellular RNAs from the same cultures. The Packaging Coefficients were then calculated in the usual way and are shown in Figure 5.2.

The C^{pac} values for the "assay" RNAs corresponding to the insertion of the large subfragments, 11a (RNA nucleotides, 0-1,465), 458 (3,063-5,324) and 459 (1,461-3,067), into pMA924 are very similar and all represent a certain degree of specific packaging. These values are somewhat less than those obtained in Chapter 4; this may reflect the smaller contributions of the Ty homologous RNA towards the "assay" RNAs. These observations

Figure 5.2. The trans-packaging of subgenomic Ty fragments. The autoradiographs correspond to the Northern filters that were used for the determination of the Packaging Coefficients of the "assay" RNAs from PT924-11a, PT924-458, PT924-459, PT924-452, PT924-453, PT924-454 and PT924-457. The C^{pac} values obtained for each "assay" RNA are shown beneath the autoradiographs. The tracks labelled 'C' contain total cellular RNA, whilst those labelled 'V' contain the RNA from the Ty-VLP containing fractions. The hybridisation probes were the 11 and IFN-1 fragments. Beneath the C^{pac} values is the genomic organisation of Ty1-15 along with the origin of the subgenomic Ty fragments that were inserted into pMA924; these contain the Ty sequences between the following restriction sites (the coordinates of the nucleotides, with reference to the 5.7 kb Ty RNA (see Figure 4.8) that these fragments span, are indicated in brackets). 11a, XhoI in 5' LTR and 5' BglII (Ø-1,467); 458, middle and 3' BglIIs (3,Ø63-5,324); 459, 5' and middle BglIIs(1,461-3,Ø67); 452, PstI and 5' BglII (896-1,465); 453, PstI and deletion endpoint of pKT357 (896-1,196); 457, HpaI and PstI (576-9Ø2); 454, PvuII and HpaI (236-576).


Trans-packaging of subgenomic Ty fragments
PT 924-11a
PT 924-458
PT 924-459
PT 924-11a
PT 924-452
PT 924-453
PT 924-454
PT 924-457
C V C V C V C V C V C V C V C V
'helper'
'assay'
0·5 0·6 0·6 Packaging Coefficient 0·5 0·4 0·1 0·1 0·2
X PvHp Pst Bg Bg Bg X
TYA
TYB
11a
458
459
452
453
454
457

imply that there are regions throughout Ty1-15 that, when transcribed into RNA, are able to direct the specific packaging of that RNA into Ty-VLPs.

The analysis of the smaller fragments (452, 453, 454 and 457 in pMA924) was aimed at localising the regions that are responsible for packaging within RNA that corresponds to the 11 fragment. The fragments 453 (899-1,196), 454 (236-576) and 457 (576-899) show no evidence of directing the packaging of RNA, but the 452 fragment (899-1,465), that is homologous to the region between the PstI and BglII sites of the 11 fragment, has a C^{pac} of Ø.4 and therefore does direct RNA packaging. This effect is lost with the deletion, from the 452 fragment, of the 269 bp between the BglII site and the pKT357 endpoint (the remaining fragment is designated 453). This result suggests that these 269 nucleotides are important for RNA packaging, and is therefore in contrast to the dramatic increase of C^{pac} that was observed when these same nucleotides were deleted in the 11 fragment deletion series (PT227-1 compared to PT357-1, Section 4.6). It can not be ruled out that the inability of the 453, 454 and 457 fragments, when inserted into pMA924, to package their corresponding "assay" RNAs is a consequence of either their requirement for interactions with RNA sequences that are elsewhere (for example, within the first 236 nucleotides of the 5.7 kb Ty RNA) or their smaller size, leading to the "shielding" of packaging signals by the remainder of the RNA molecule. Therefore, the use of pMA924 for the analysis of Ty RNA packaging signals is straightfoward and has produced results, so far, that are relatively easy to interpret. However, it is important to realise that the analysis

of Ty RNA that is flanked on both sides by heterologous sequences may not be wholely representative of the *in vivo* situation.

5.4 Conclusions and Discussion

It was previously shown that RNA sequences homologous to the regions 5' to the PvuII site of *TYA*, up to the point of transcription initiation, and 3' to the BglII site (at the 3' end of *TYB*) up to the point of transcription termination are both involved in the packaging of Tyl-15 RNA into Ty-VLPs. The results shown in Figure 5.2 clearly show that the three large subfragments of Tyl-15 that cover the whole of the unique region (epsilon), when transcribed, also direct the specific packaging of Ty RNA into Ty-VLPs. It was further implied that the regions within the 11 fragment RNA that are responsible for this reside 3' to the site corresponding to the PstI site of *TYA* (RNA nucleotide, 896). This is in agreement with the suggestion that signals downstream from the site equivalent to the pKT359 endpoint (RNA nucleotide, 737) are important for RNA packaging (Section 4.6).

Therefore Ty RNA, like the genomic RNAs of retroviruses, has sequences that direct its specific packaging into viral particles, in this case Ty-VLPs. The apparent dispersal of these sequences throughout the Ty RNA, and therefore throughout the genome of Tyl-15, is very different to the localised packaging signals of the retroviral RNAs. However, there is a major difference between the biological requirements of the two

systems. Retroviral packaging sequences are restricted to regions that are present in the genomic RNA but not in its spliced product, the env-mRNA (Section 4.1). This ensures that the genomic RNA is packaged into virions whilst the env-mRNA is not. There is no such constraint for the position of the packaging signals within Ty RNA because there is no requirement for the discrimination between the packaging of different Ty RNA species. If the Ty-VLPs are intermediary in Ty element transposition (Section 3.5), the additional packaging of the smaller 5.0 kb and 2.2 kb Ty RNA species, by virtue of the dispersed packaging signals, will not result in their insertion as "mutated" Ty elements into the yeast genome. Unlike the env-mRNA of retroviruses, these RNAs are not produced by the splicing of the 5.7 kb Ty RNA, and therefore do not possess the same 5' and 3' termini as the genomic RNA. This is likely to result in their inability to be reverse transcribed and integrated. For example, the 5.0 kb RNA (SPT3 RNA), which initiates 800 bp downstream from the 5.7 kb RNA start site, does not possess the sequences for the binding of the proposed reverse transcriptase primer, $tRNA^{met}$.

The finding that there is not a localised region within Ty1-15 that is responsible for the packaging of Ty RNA into Ty-VLPs was unexpected in the light of the information regarding the packaging of retroviral RNAs. Therefore there was no simple deletion mutation, analogous to the ψ^- mutation of MoMLV, that could be made to largely prevent both RNA packaging and, by supposition, Ty element transposition. It is predicted that small deletions within Ty1-15 should reduce, but not abolish, RNA

packaging into Ty-VLPs. Larger deletions would be expected to result in less packaging and therefore less transposition, provided that the processes of reverse transcription and integration were not inhibited in any way.

CHAPTER 6. TRUNCATED p1 PROTEINS AND p1-INTERFERON FUSION PROTEINS CAN BOTH ASSEMBLE INTO VIRUS-LIKE PARTICLES

6.1 Introduction

It was shown in Chapter 3 that the primary translation product of the TYA gene, the 50 kD protein p1, is able to form Ty-VLPs. The major protein of Ty-VLPs produced by the expression of a whole Ty element (TYA plus TYB) is the 48 kD protein p2; this protein is a proteolytic cleavage product of p1, as are the smaller Ty-VLP proteins p4, p5 and p6 (Adams et al., 1987). It was suggested at the end of Chapter 4 that proteins expressed from TYA genes, shortened from the 3' end, may still be able to aggregate into particles.

Therefore, a series of experiments aimed at determining the minimum protein requirements for VLP formation were carried out. Given that p1 forms VLPs, a series of plasmids were constructed that would express p1 proteins truncated by variable amounts at their 3' (or carboxyl) termini. The ability of such proteins to form particles was assayed by fractionating cell extracts on 15-45% (w/v) SW27 sucrose gradients and analysing the samples on SDS-polyacrylamide gels. Only proteins that are assembled into particulate matter will migrate beyond the top 6 or 7 fractions, rather like the peaks of p1 and p2 in Figure 3.5 (T91-10, Ty-VLPs) and p1 in Figure 3.8 (T91-11, Ty-VLPs).

6.2 Plasmid Constructions

The deletions from the 3' terminus of the TYA gene of Ty1-15 that were generated in Section 4.5, were used for this analysis. The plasmids pKT357, pKT358, pKT359, pKT450, pKT451 and pKT202,

however, use the authentic Ty expression signals, some of which are within the TYA coding region (Section 1.4.3 and Figure 4.11). To ensure that the expression levels of the different truncated proteins were compatible, the powerful PGK promoter was used for driving transcription, thus the detectibility of any particulate proteins should be equal in each experiment. The use of the PGK promoter in the pMA91 configuration also means that the RNAs corresponding to the truncated TYA genes will be efficiently translated (Section 1.4.5). Two sets of plasmids were constructed, those that would express truncated TYA genes, and those that would express truncated TYA genes fused, in-frame, to IFN coding sequences (as in Chapter 4).

The scheme for these plasmid constructions is outlined in Figure 6.1. The first step was the ligation of three PstI fragments so that the first TYA deletion endpoint (pKT357) was brought under the transcriptional control of the PGK promoter. Figure 6.2A shows the colony hybridisation that was necessary for the identification of Amp^R recombinants containing the IFN sequences. Figure 6.2B shows PstI digested minipreps of the IFN-containing recombinants, lane 1 contains the three PstI fragments in equimolar amounts: this plasmid was designated pMA91-357-1. The IFN and Ty flanking sequences were removed from pMA91-357-1 and replaced with a 300 bp SalI-BamHI modified PGK 3' transcription terminator fragment that contains translation stop codons in all three reading frames. The resulting plasmid was designated pMA5620 (it will be discussed further in Chapter 7). The other deletion endpoints within TYA were inserted into

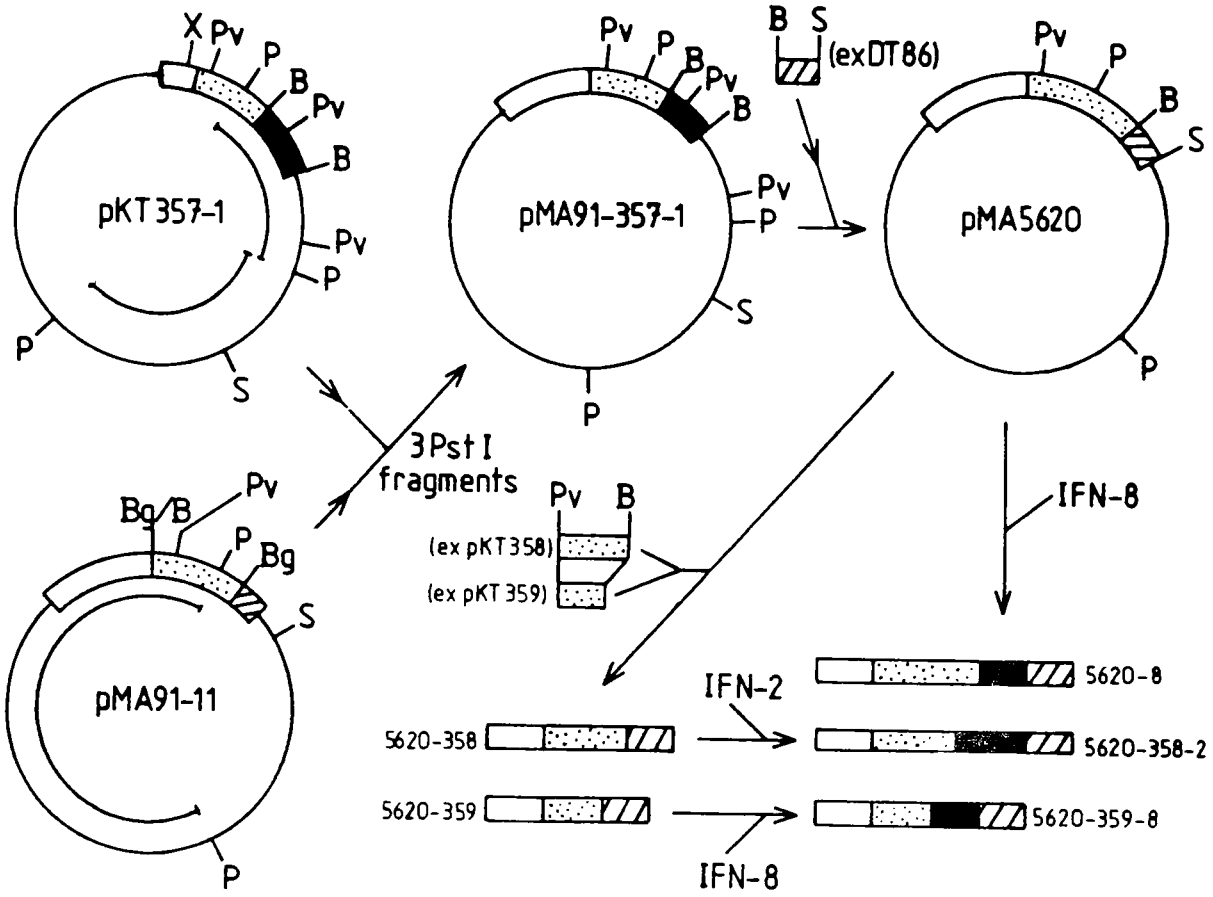


Figure 6.1. The construction of the plasmids containing PGK driven truncated TYA genes and truncated TYA genes fused to the α2-interferon gene. Thin lines, yeast/E. coli selection and replication sequences (similar to those of pMA91, Figure 1.8) and unspecified yeast sequences flanking Tyl-15 (similar to those of pKT227 (Figure 4.6), pKT357-1 and pMA91-357-1 only); open boxes, promoter sequences (the 5' LTR of Tyl-15 and the 5' region of the PGK gene); stippled boxes, IFN sequences; hatched boxes, terminator sequences from the 3' region of the PGK gene; X, XhoI; Pv, PvuII; P, PstI; B, BamHI; S, SalI; Bg, BglII. The three PstI fragments used for the construction of pMA91-357-1 are indicated.

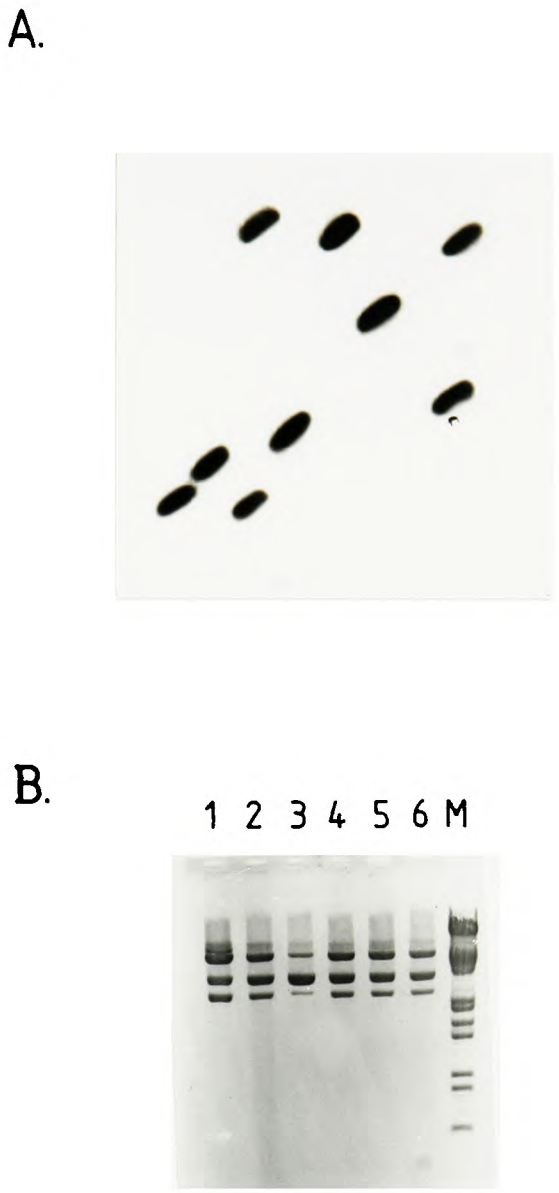


Figure 6.2. A, autoradiograph of the colony hybridisation used to ientify the IFN containing recombinants in the construction of pMA91-357-1; the hybridisation probe was the IFN-1 fragment. B, 1% Tris-borate agarose minigel showing the PstI digestion of miniprep DNA prepared from six of the interferon containing clones identified in A; the markers (M) were described in the legend to Figure 4.7.

pMA5620 by straightforward PvuII-BamHI fragment replacements to produce the plasmids, pMA5620-358, pMA5620-359, etc. The IFN-2 or IFN-8 fragments were inserted into the unique BamHI sites of these plasmids to produce the plasmids pMA5620-8, pMA5620-358-2, pMA5620-359-8, etc. The IFN fragment was selected for its maintenance of reading phase with TYA, as in Section 4.5. The IFN-8 fragment (same 5' terminus as IFN-1, Figure 4.3) could be used because of the presence of the PGK terminator fragment. All the plasmids were transformed into the yeast strain MD40-4c.

6.3 The Ability of Truncated p1 Proteins to Form Particles

The primary translation product of the TYA gene of Ty1-15, p1 (50 kD), is able to assemble into particles. This was detected by the migration of p1 beyond the top 6 or 7 fractions of a 15-45% (w/v) SW27 sucrose gradient fractionated T91-11 extract, as visualised on an SDS-polyacrylamide gel (Figure 3.8). Samples from similarly fractionated T5620 and T5620-358 extracts were therefore analysed in the same way. The T5620 extract has a clear peak of particulate protein of ~48 kD between fractions 8 and 13 (Figure 6.3). It is interesting that this protein peak is higher up the gradient than the VLPs in T91-10 and T91-11 extracts (Figures 3.5 and 3.8, respectively), this reflects a decrease in the particle density. The fractionated T5620-358 extract had no obvious peak of particulate protein, though there was a faint peak of a ~50 kD protein between fractions 5 and 9 (data not shown). This gel was therefore very similar to the one representing the fractionated T91 extract (Figure 3.6), with the

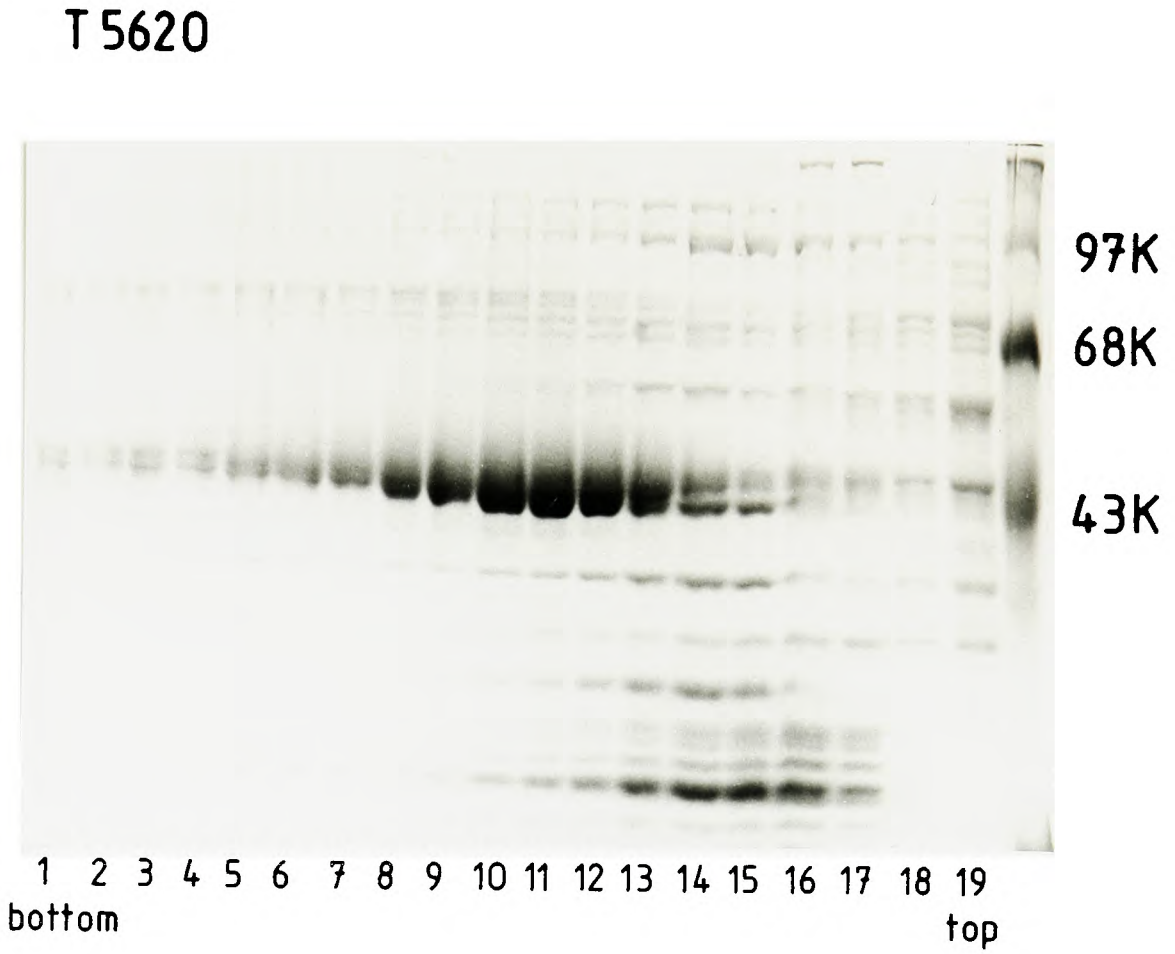
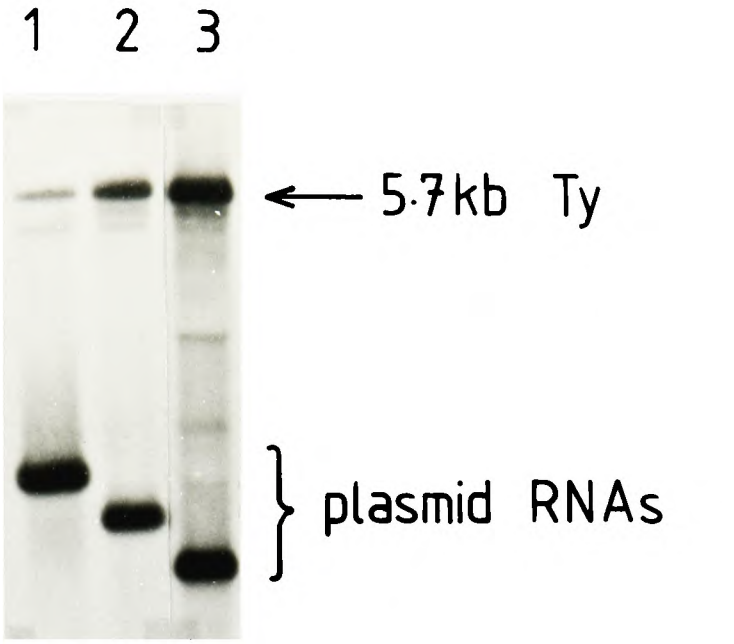
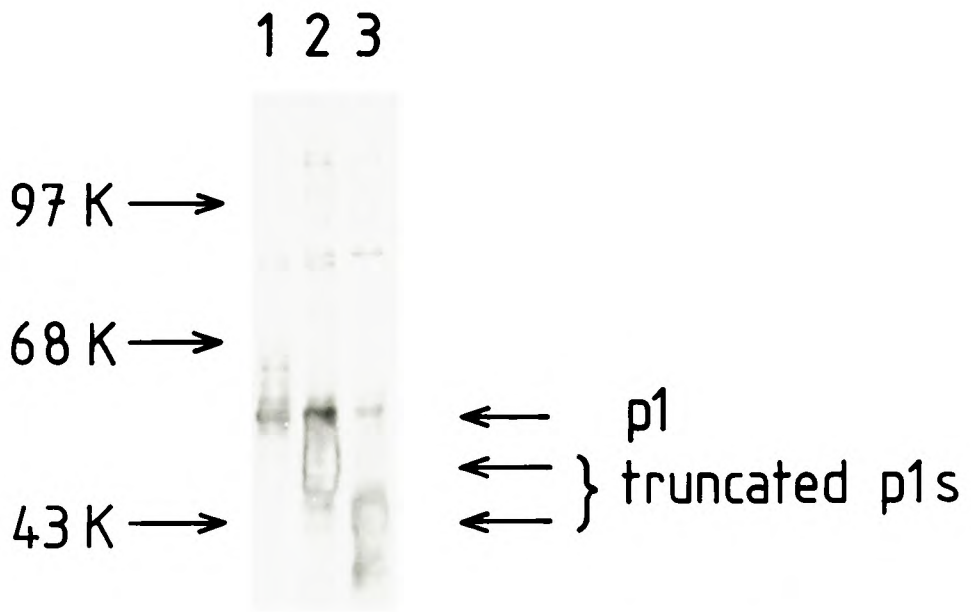


Figure 6.3. Coomassie Blue-stained gradient profile of a T5620 extract. The preparation and treatment of the samples and the gel were described in the legend to Figure 3.5.

Figure 6.4. The analyses of the total cellular RNA by Northern blotting (upper panel) and of total cytoplasmic protein by Western blotting (lower panel) of extracts from the T91-11, T562Ø and T562Ø-358 cell cultures that were used for the detection of particles (Figures 3.8, 6.3 and Section 6.3, respectively). The hybridisation probe for the Northern blot was the 11 fragment; the bands corresponding to the Ty specific RNAs from the plasmids and the endogenous Ty elements (5.7 kb) are indicated. A 12.5% (3Ø:Ø.8) SDS-polyacrylamide gel was run for the Western blot analysis and an anti Ty-VLP antiserum (Adams *et al.*, 1987) was used for the hybridisation; the p1 present in the T91-11 extracts and truncated pIs present in the T562Ø and T562Ø-358 extracts are indicated.





1. pMA 91-11
2. pMA 5620
3. pMA 5620-358

faint peak corresponding to the endogenous Ty-VLPs.

A Northern blot analysis of the total cellular RNA and a Western blot analysis of the total cytoplasmic protein of T562Ø and T562Ø-358 were carried out (Figure 6.4). T91-11 was assayed as a positive control for both analyses. This data confirms that T562Ø and T562Ø-358 both contain adequate RNA and protein, corresponding to the truncated TYA genes, for the detection of particles by the above approach. Thus, the failure to detect a peak of particulate truncated p1 in the T562Ø-358 extract was not a consequence of the protein being in insufficient levels.

Therefore, a p1 protein that has been truncated by 6Ø amino acids at its C-terminus, so that only 381 codons of TYA remain (pMA562Ø), is still able to assemble into particles. Whilst, the removal of a further 95 codons from TYA so that only 286 amino acids of p1 remain (pMA562Ø-358) does, however, prevent the formation of particles.

6.4 The Ability of p1-IFN Fusion Proteins to Form Particles

There were two reasons for assaying the ability of the p1-IFN fusion proteins to form particles. Firstly, these fusion proteins are the same as those that were present in the packaging assay transformants in Section 4.6; therefore the ability of those various "assay" RNAs to be translated into proteins capable for forming particles can be determined. This information might allow a more detailed interpretation of the packaging results that were presented in Section 4.6. Secondly, the inability of

the truncated p1 protein in the T5620-358 extract to form particles might have been a consequence of the protein having insufficient mass to form particles. The addition of a further 20 kD of IFN protein would increase the mass so that particle formation might ensue.

The p1-IFN fusion proteins present in the extracts from the two strains T5620-8 and T5620-358-2 were therefore assayed for their ability to form particles as before. A sucrose gradient fractionated T5620-8 extract was analysed on an SDS-polyacrylamide gel (Figure 6.5). There is a broad peak of a ~68 kD particulate protein between fractions 2 and 11, this is the predicted size of the p1-IFN fusion protein in T5620-8. However, when a T5620-358-2 extract was treated in a similar way, there was no such particulate peak corresponding to the p1-IFN fusion (data not shown); the gel was therefore reminiscent of those obtained with fractionated T91 and T5620-358 extracts. The RNA and protein corresponding to the TYA:IFN fusions were shown to be present at levels sufficient for the detection of particles (Figure 6.6).

Therefore, the addition of 20 kD of IFN protein onto the truncated protein with 381 amino acids of p1 remaining (encoded by pMA5620-8) does not affect its ability to form particles; it does, however, increase the sedimentation rate of the particles so that they migrate with a similar rate as T91-10 and T91-11 Ty-VLPs. The addition of IFN onto the protein with only 286 amino acids of p1 remaining (encoded by pMA5620-358-2) does not restore the ability of the protein to assemble into particles.

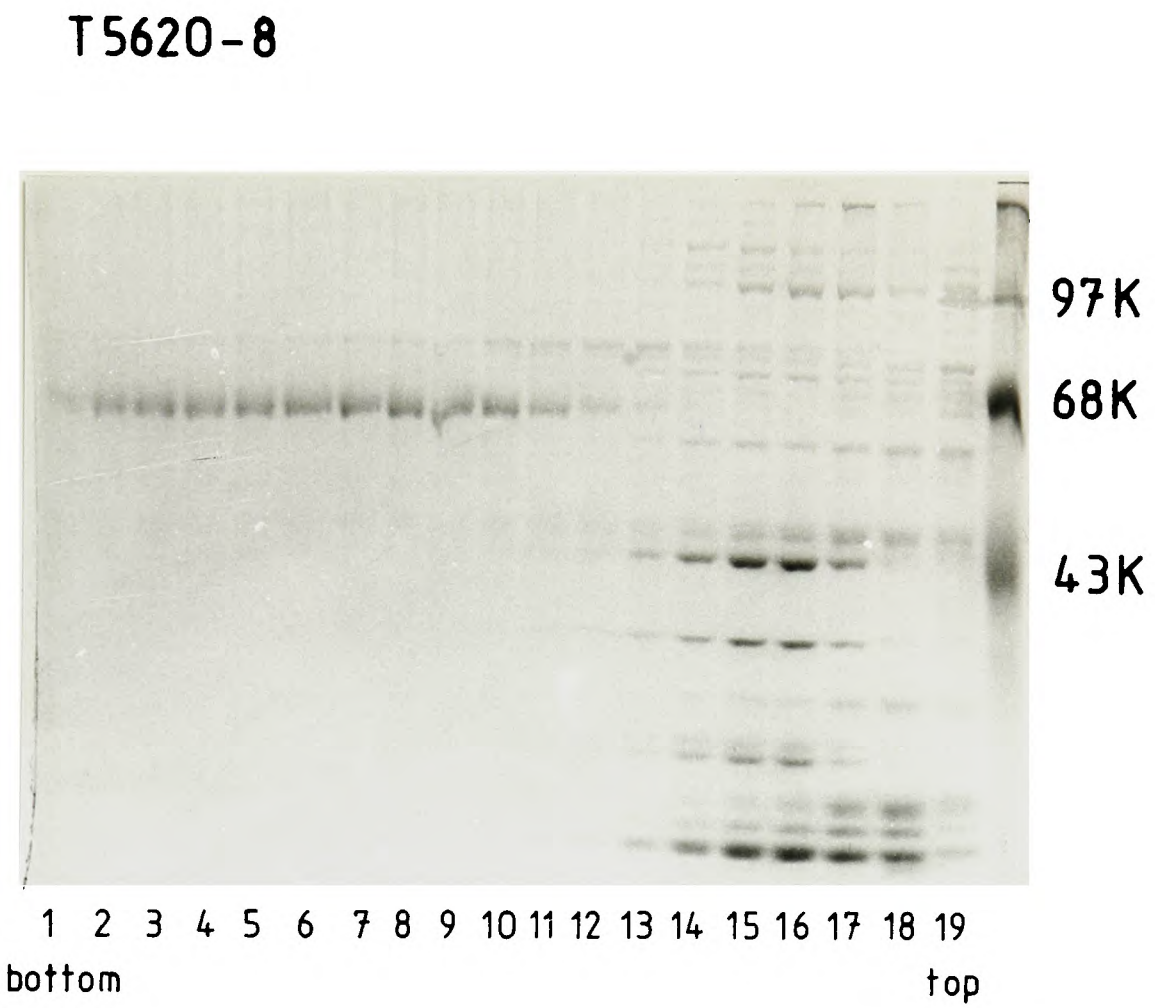


Figure 6.5. Coomassie Blue-stained gradient profile of a T5620-8 extract. The preparation and treatment of the samples and the gel were described in the legend to Figure 3.5.

6.5 Electron Microscopy and Discussion of Particle Formation

To confirm the particulate nature of the truncated p1 protein and the p1-IFN fusion protein from the T562Ø and T562Ø-8 extracts, respectively, the peak fractions were pooled, pelleted and examined in the electron microscope (Figures 6.7 and 6.8). The T562Ø particles are ~6Ø nm in diameter and have an electron dense core with a lighter staining shell (Figure 6.7). This morphology is surprisingly similar to that of the T91-1Ø Ty-VLPs, though their sedimentation rates are rather different (Figures 3.4 and 6.3). The major protein of the T91-1Ø Ty-VLPs is p2 and it is interesting to speculate that the 3' shortening of TYA by 6Ø codons in pMA562Ø might result in the expression of a truncated p1 that is very similar to p2. This suggests that the in vivo proteolytic cleavage of p1 to yield p2 might be mimicked by the engineered truncation of p1 during the expression of the 3' deleted TYA gene in pMA562Ø. The T562Ø-8 particles, known as hybrid particles, retain the usual diameter for Ty-VLPs of ~6Ø nm, but possess a morphology that is radically different from that of any other Ty related VLPs (Figure 6.8). They are "woolly" in appearance, and this can be attributed to differential staining with uranyl acetate, as a result of their altered structure.

The deletion present in pMA562Ø-358 that removes 155 codons from the 3' end of TYA is sufficient to prevent the corresponding truncated p1 protein from forming particles. This implies that enough of the amino acids that are involved in the protein-protein interactions of particle formation have been deleted so

Figure 6.7. Electron micrograph of uranyl acetate-stained Ty-VLPs from the pelleted fractions of a SW27 sucrose gradient fractionated T562Ø extract (X 3Ø,16Ø).

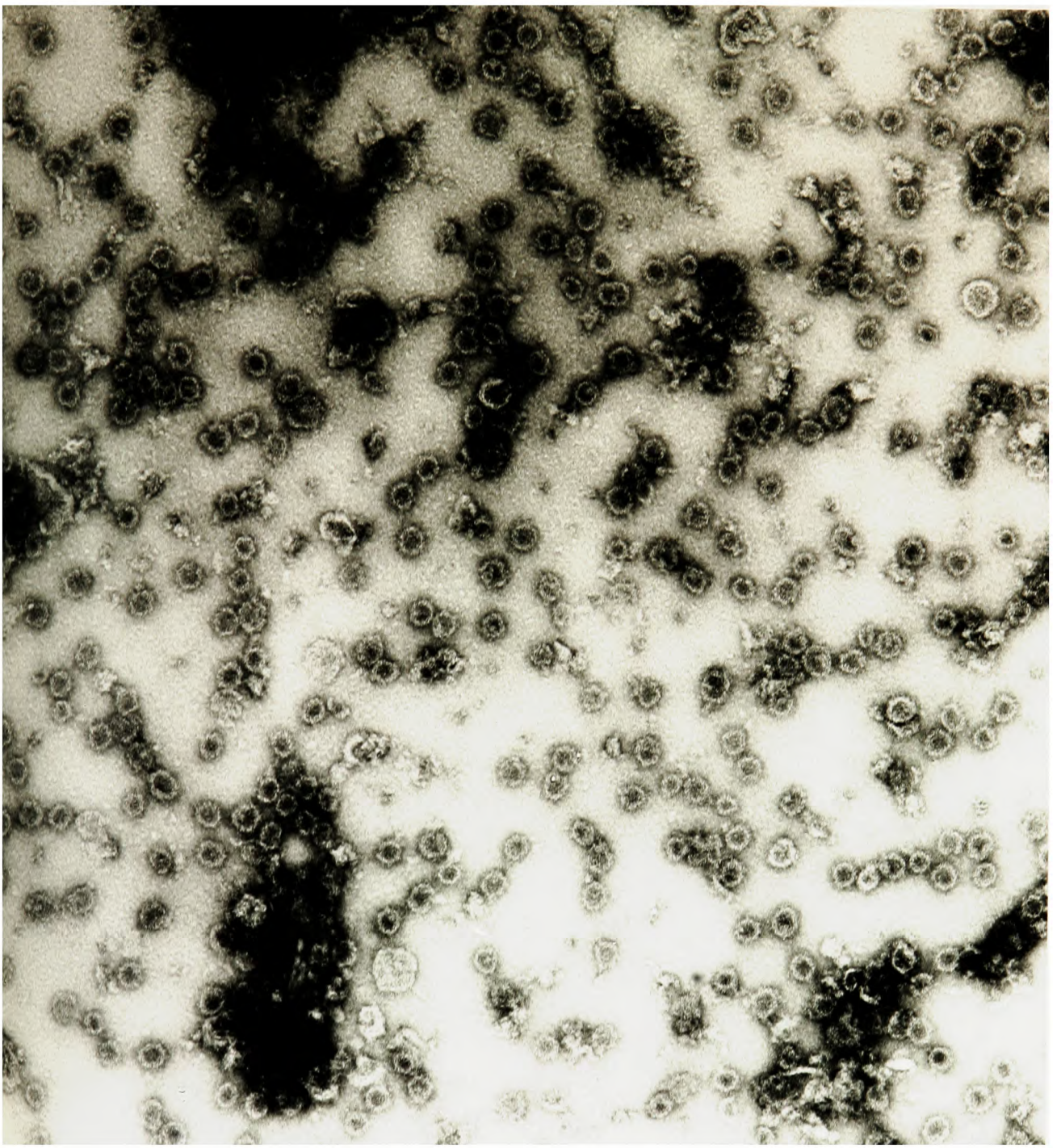
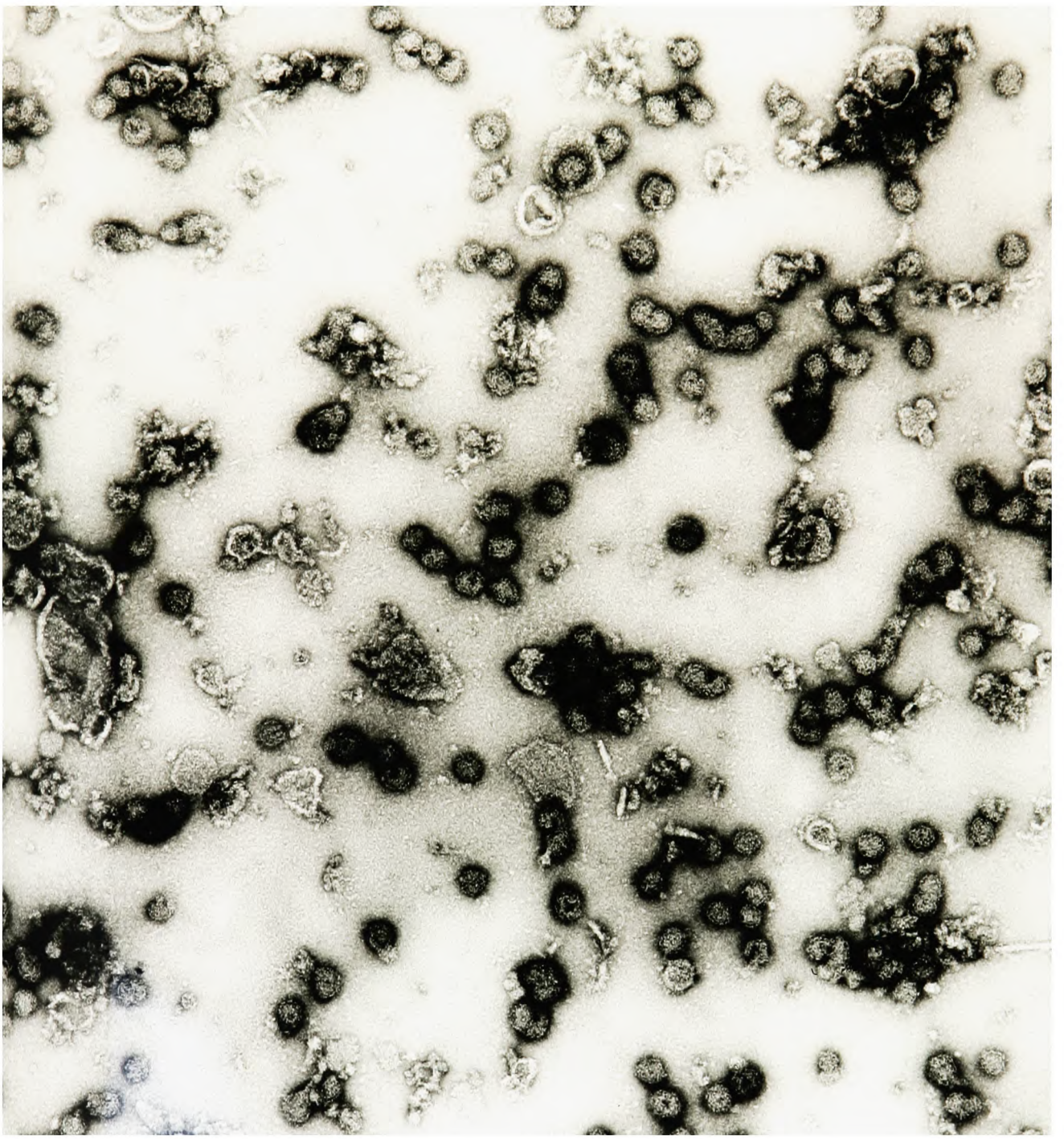


Figure 6.8. Electron micrograph of uranyl acetate-stained hybrid Ty:IFN-VLPs from the pelleted fractions of a SW27 sucrose gradient fractionated T5620-8 extract (X 30,160).



as to prevent assembly. The amino acids that are involved in such interactions may include the numerous proline residues situated towards the carboxyl terminus of p1. The requirements for RNA-protein interactions in VLP formation are currently unknown; retroviral particles that lack RNA have been identified, so it is possible that these interactions are not pre-requisites for particle formation. There are a large number of amino acid residues with positively charged side chains (arginine and lysine) towards the amino terminus of p1, and these have been predicted as being instrumental in protein association to the negatively charged RNA (Adams et al., 1987). The effect on VLP formation of deleting such residues from p1 whilst leaving the carboxyl terminus intact has not been studied to date. It is also known that the VLPs present in the T91-11-8 and T562Ø-8 strains both contain some IFN homologous RNA (data not shown).

The results described in this chapter highlight the problems involved with the interpretation of the packaging assay results shown in Chapter 4, and thereby justify, to some extent, the application of pMA924 for assessing packaging signals (Chapter 5). For example, PT357-1 cells contain both mature Ty-VLPs (produced by pMA132-1Ø) and hybrid Ty-VLPs (produced by pKT357-1). The cofractionation of these particles (data not shown) means that the actual protein, and therefore VLPs, to which an RNA molecule is associated can not be determined easily. The C^{pac} of the "assay" RNA of PT357-1 was the highest obtained for any of the RNAs tested (Figure 4.11). Efficient packaging of the "assay" RNA into the hybrid particles could be responsible for this,

though such an explanation seems unlikely for two reasons. Firstly, the VLPs in T562Ø-8 display only low levels of RNA packaging (data not shown), and secondly, it has been shown that the variable levels of the additional particle forming protein, p1, in PT91-11-8, PT346-8 and PT219-8 did not seem to influence the C^{pac} values obtained for their "assay" RNAs (Figure 4.5). Thus, the high C^{pac} values observed for PT357-1 are more likely to be the result of the "assay" RNA forming a secondary structure that is favourably packaged. Alternatively, the presence of the ~68 kD particulate fusion protein may have some effect.

CHAPTER 7. HYBRID TY VIRUS-LIKE PARTICLES CAN FUNCTION AS SELF-ASSEMBLING EPITOPE CARRIERS

7.1 Introduction: Recombinant DNA Approaches to Vaccine Development

The ability of a p1-interferon fusion protein to aggregate into particles has been demonstrated by the protein's sedimentation rate through a sucrose gradient and the subsequent examination of the "peak" fractions in an electron microscope (Chapter 6). This raised the possibility that hybrid Ty-VLPs might represent a novel method of purifying and presenting antigens, and might therefore function as self-assembling epitope carriers. Self-assembling subunit antigens constitute one of several modern approaches to vaccine development that have been made possible by the advent of recombinant DNA technology.

Vaccines have been in existance for many years. The three classical types are live, killed and subunit vaccines. These have been highly successful in some instances, for example the worldwide use of vaccinia virus vaccines led to the erradication of smallpox by the late 1970's (Brown et al., 1986). The application of live (or attenuated) vaccines like this carry the risk of virus mutation; this may result in the vaccine strain becoming virulent and therefore dangerous in itself. The killed and subunit vaccines do not carry this particular risk, but the preparative procedure by which they are made often reduces the immunogenicity of the antigenic determinants, thus weakening the immune response. In addition, the handling of large quantities of the infectious agent during the manufacturing process can be dangerous, this makes the vaccine rather expensive to produce and its application may therefore be uneconomical (reviewed by

Shinnick et al., 1983).

Two of the approaches that are being tested in attempts to overcome the technical and economic limitations of some of the conventional vaccines involve the use of recombinant DNA techniques and the chemical synthesis of peptides. Synthetic peptides consisting of a few amino acids have been shown to elicit antbodies that can react with the relevant full length protein (Shinnick et al., 1983). However, the identification of the best peptide and the selection of the proper adjuvant for antigen presentation both represent considerable problems for this approach. Also, there is scant evidence to suggest that synthetic peptides are capable of evoking cell-mediated responses. For a vaccine to be effective it needs to induce both humoral (antibody) and cytotoxic (cell-mediated) immunity to antigens. The success of the live, smallpox vaccine is partly due to its induction of both these responses; thus an ideal antigen presentation system should be "viral-like" in its nature. It should have a high molecular weight and be polymeric with multiple antigenic determinants that are displayed in an effective conformation. The recombinant vaccines that are being tested have been designed with these criteria in mind.

Recombinant subunit vaccines involve the expression of the gene that codes for the neutralising antigen in a suitable host-vector system. Efforts to produce a vaccine against Hepatitis B Virus (HBV) are being concentrated on this approach. The expression of the mature surface antigen (HBsAg) in yeast results in its self-assembly into 22 nm particles (Valenzuela et al.,

1982). These can be easily purified and the vaccination of chimpanzees with them provides complete protection against infection when challenged with live HBV (Murray et al., 1984; McAleer et al., 1984). Similar HBsAg particles have been produced and secreted from rodent cells (Patzer et al., 1986). These are more immunogenic and this may reflect the increasingly authentic glycosylation occurring in mammalian cells. Attempts to produce vaccines for other diseases, by this approach, have been less successful because the expressed antigens are often unable to self-assemble.

The expression of a pathogen derived gene within a hybrid virus so that the neutralising antigen is presented to an immune system by the virus, rather than the pathogen, would satisfy the criteria for good antigen presentation without causing the disease. Live recombinant virus vaccines are being developed along these lines; most studies have utilised vaccinia virus because it has the advantages of being well understood, having a broad host range and possessing a large genome that can be conveniently manipulated (Mackett et al., 1982; Panicali and Paoletti, 1982; Brown et al., 1986). The gene of interest is engineered into a non-essential part of the vaccinia genome (usually the thymidine kinase gene), this recombinant virus is propagated and can then be administered dermally. The limited replication of the virus that follows, can result in the induction of both humoral and cytotoxic responses to a variety of pathogens. Using this system, possible live vaccines against the AIDS virus (HIV) (Chakrabarti et al., 1986; Kieny et al., 1986; Zarling et al., 1986), influenza (Bennink et al., 1984),

hepatitis (Moss et al., 1984; Paoletti et al., 1984), rabies (Kieny et al, 1984) and malaria (Smith et al., 1984) are being tested and have generated much interest. However, major concerns over the broadening of the host range of vaccinia and viral mutation to create novel pathogens do exist (Brown et al., 1986).

The advent of a third type of recombinant vaccine, the self-assembling epitope carriers, combines the attractive features of the two previously described approaches to vaccine development. The production of a hybrid protein, by gene fusion, that comprises at least two functional domains, one which causes protein aggregation and one which corresponds to the epitope (or neutralising antigen), has the advantages of being safe (unlike the live, vaccinia based vaccines) and reliably self-assembling (unlike the recombinant subunits) as well as satisfying the previous criteria for effective antigen presentation. Some early success has been achieved with TMVCP-polio (Tobacco Mosaic Virus Coat Protein), HBsAg-HSVlgD (Herpes Simplex Virus) and HBsAg-polio fusion proteins (Haynes et al., 1986; Valenzuela et al., 1985; Delpeyroux et al., 1986) but problems with low yields, poor subunit aggregation and immunodominance have been encountered. The ability of hybrid Ty-VLPs to act as self-assembling epitope carriers was therefore tested: this system is particularly attractive because high yields of relatively pure Ty-VLPs can be extracted easily from yeast.

7.2 The TYA:IFN Fusion Particles have Ty and Interferon Epitopes

The p1-IFN fusion protein that was identified by its particulate nature in T5620-8 extracts was ~68 kD when analysed by SDS-polyacrylamide gel electrophoresis (Figure 6.5). This was the predicted mass of a fusion protein produced by the in-frame fusion of a TYA gene, truncated by 60 codons at its 3' end, and the α2-interferon cDNA fragment (IFN-8). Figure 7.1 shows the structure of pMA5629-8 and its construction was described in Section 6.2. To confirm that the ~68 kD protein was the predicted p1-IFN fusion, it was tested in a Western blotting experiment. Duplicate SDS-polyacrylamide gels of sucrose gradient fractionated T5620-8 extracts were run as before (Figure 6.5). The separated proteins of each gel were transferred to nitrocellulose and hybridised with either an anti Ty-VLP antiserum (Adams et al., 1987) or an anti interferon antibody for the first reaction (Figure 7.2, a and b respectively). Both profiles have a clear peak of hybridisation between fractions 3 and 10 at a mobility corresponding to ~68 kD. This coincides with the profile of the previously identified p1-IFN fusion protein (Figure 6.5) and demonstrates that the ~68 kD protein carries both Ty and IFN epitopes. The Western profile obtained with the anti Ty-VLP antiserum also has a minor, co-sedimenting peak of a ~48 kD protein, this is presumably the p2 of the Ty-VLPs that are produced in low levels by the 30-35 endogenous Ty elements. The novel morphology of T5620-8 hybrid VLPs was described and discussed earlier (Section 6.5).

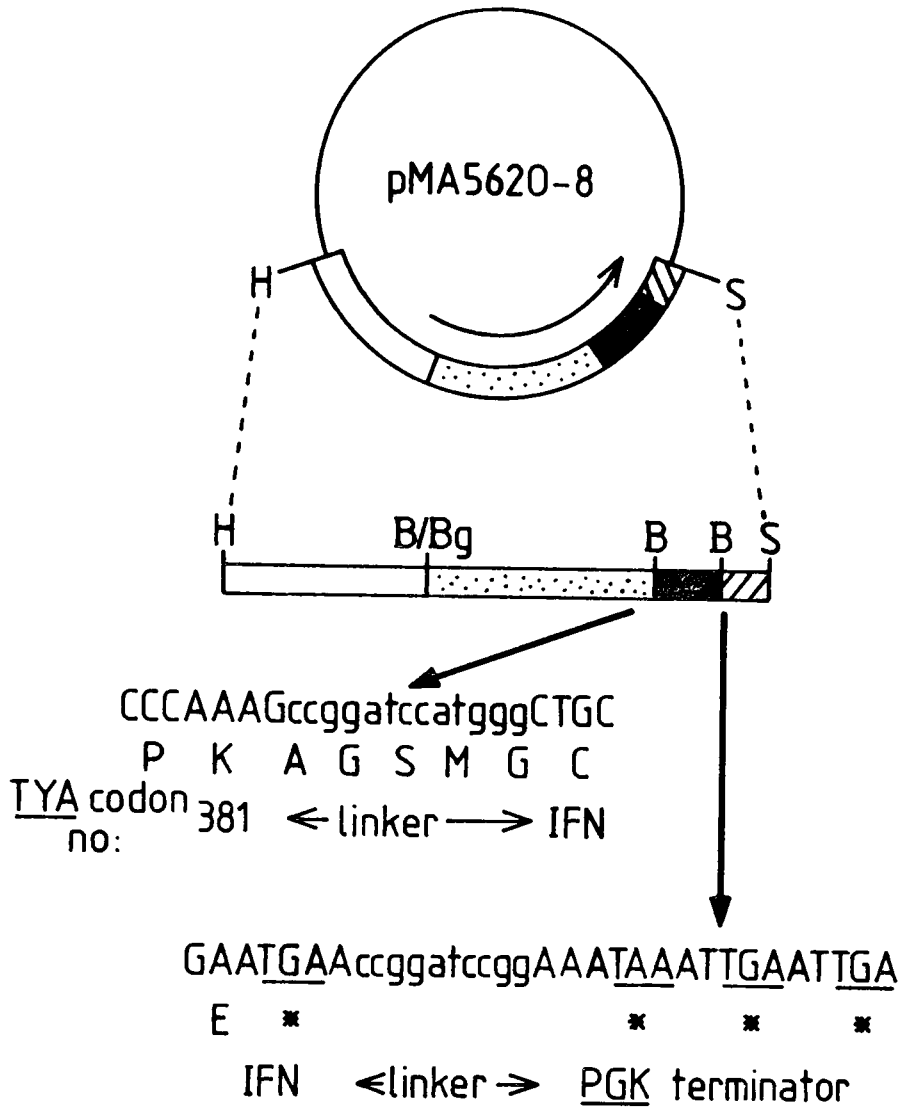


Figure 7.1. The structure of pMA5620-8 and the nucleotide sequences and deduced amino acid sequences of the TYA:BamHI linker:IFN and IFN:BamHI linker:PGK terminator junctions. Thin line, yeast/E coli selection and replication sequences (similar to those of pMA91, Figure 1.8); open box, 5' promoter region from the yeast PGK gene; stippled box, truncated TYA gene; solid box, α2-IFN sequences; hatched box, a modified 3' transcription terminator from the yeast PGK gene, with stop codons in all three reading phases; the arrow indicates the direction of transcription promoted by the PGK 5' region; H, HindIII; Bg, BglII; B, BamHI; S, SalI; upper case nucleotides are derived from the TYA, IFN or PGK sequences; lower case nucleotides are derived from the linker sequences; *, translation stop codons.

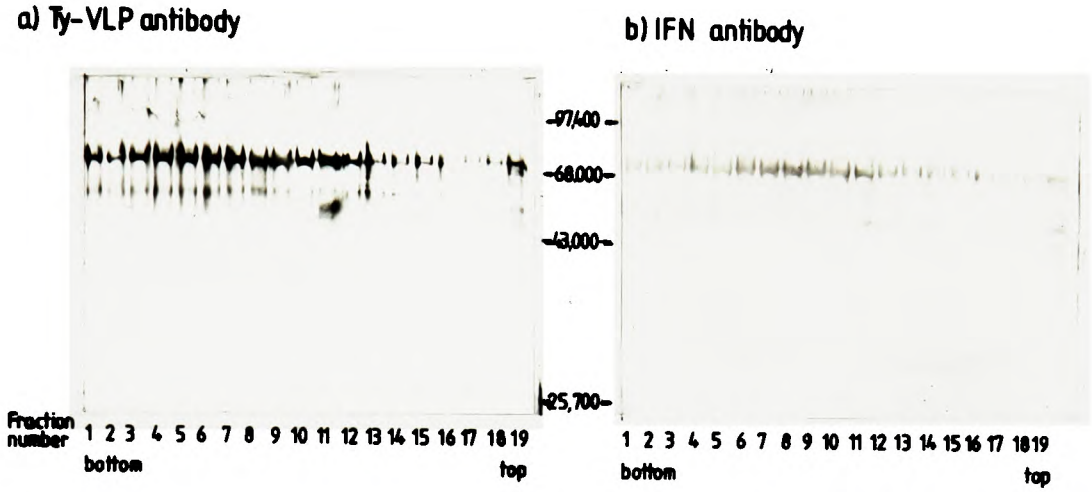


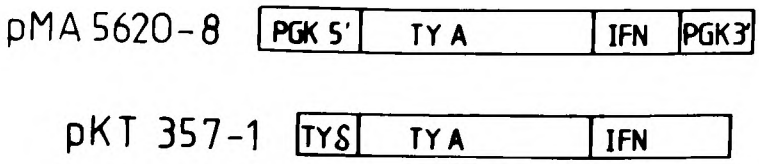
Figure 7.2. Western blot analyses of sucrose gradient profiles of a T562Ø-8 extract (Figure 6.5). The proteins were separated on 9% (3Ø:1.6) SDS-polyacrylamide gels and the filters were probed with either an anti Ty-VLP antiserum (a) or an anti interferon antibody (b).

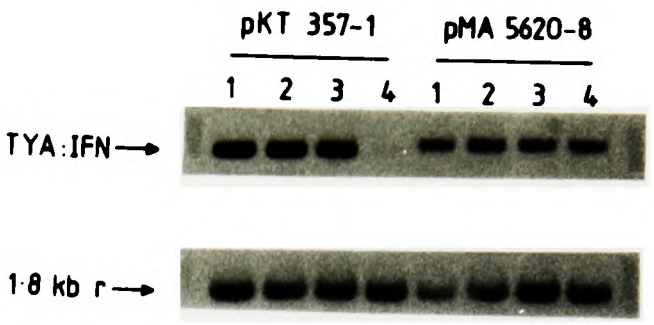
7.3 The Ty:IFN Hybrid VLPs Elicit an Antibody Response Against Interferon

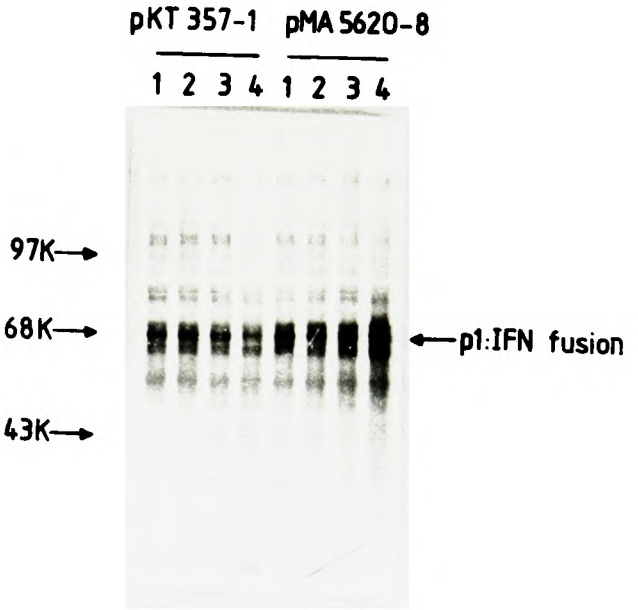
Since the pl-IFN hybrid VLPs display both Ty and IFN epitopes, their efficacy in eliciting an antibody response to their IFN component was investigated. Such a response would resemble those observed for the self-assembling epitope carriers that were described earlier (Section 7.1).

Before this was done, the parameters required for maximising the yields of the pl-IFN fusion protein were characterised. Both T562Ø-8 (PGK directed expression) and T357-1 (Ty directed expression) cultures were analysed for their TYA:IFN fusion RNA and protein contents at a range of cell densities (Figure 7.3). Both plasmids have the same TYA:IFN ORF and will therefore produce the same fusion protein. The upper panels show the RNA levels at the various densities: the lower of the two Northern blot analyses demonstrates that equal amounts of RNA were loaded in each track. The levels of the fusion RNA are 2-5 fold higher when transcription is driven by the authentic Ty expression signals, however when the cells are in stationary phase, the Ty transcription is abolished. The transcriptional inactivity of wild-type Ty elements (using the authentic Ty expression signals) in cells that are in stationary phase has been previously demonstrated in rich media (Elder et al., 198Ø) but not in selective media (as it is here). The levels of the fusion protein do not reflect their respective RNA levels (lowest panel). At all cell densities, the protein levels are higher in the T562Ø-8 cells. This shows once again that the PGK promoter,

Figure 7.3. The analyses of the total cellular RNA by Northern blotting (upper panels) and total cytoplasmic proteins by Western blotting (lowest panel) of the pl-IFN fusion protein containing T5620-8 and T357-1 extracts at various cell densities (as indicated). The hybridisation probes for the Northern blot filter were the 11 fragment (specific for the TYA:IFN hybrid RNA) and the 1.8 kb EcoRI fragment from pY1rG12 (Petes et al., 1978) (specific for the 1.8 kb ribosomal RNA, this demonstrates the equal loading of RNA in each track). The proteins for the Western blot were separated on a 12.5% (30:0.8) SDS-polyacrylamide gel and the filter was probed with an anti Ty-VLP antiserum; the position of the pl-IFN fusion protein is indicated.







1. 4×10^6 cells / ml
2. 8×10^6 cells / ml
3. 2×10^7 cells / ml
4. Stationary (2×10^7 plus 24hrs)

Growth in 1% Glucose

in the pMA91 configuration, results in the increased translation efficiency of a heterologous gene; in this case the lower RNA levels are more than compensated for. The levels of the fusion protein are at their highest in T562Ø-8 extracts from cells that have reached stationary phase. The half life of the protein appears to be extended due to its particulate nature, thus causing its accumulation. This observation of a particulate protein in yeast having a prolonged half life is not without precedent, for example, aggregated radiolabelled hepatitis B core antigen was readily detectable after 48 hours of a pulse-chase experiment (D.L. Montgomery, personal communication). The level of the fusion protein in the stationary phase T357-1 extract is probably reduced as a consequence of the loss of transcription.

This data indicates that the T562Ø-8 extracts, rather than the T357-1 extracts, contain higher levels of the p1-IFN fusion protein and that the protein yields do not diminish in cultures with high cell densities. However, the growth of cells into stationary phase is likely to be counter-productive, due to the inability of glass beading to rupture the cells as effectively as when they are in logarithmic phase.

Hybrid VLPs were therefore prepared from a culture of T562Ø-8 that had been grown to 2 x $1Ø^7$ cells/ml, by the standard procedures that have been used previously (Section 2.14) (Figure 6.8). A relatively pure preparation of the ~68 kD fusion protein was obtained (Figure 7.4) and this was used to immunise rabbits (Section 2.16.7). The minor contamination of the fusion protein with the proteins of the endogenous Ty-VLPs is not a problem

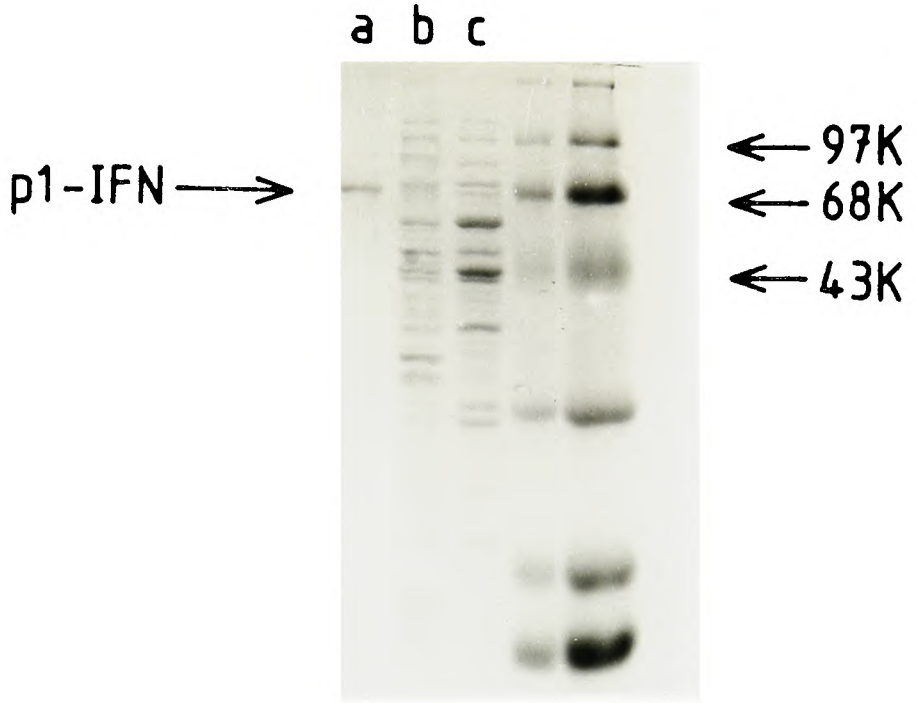


Figure 7.4. Coomassie Blue-stained gel showing the purification of the ~68 kD p1-IFN fusion protein. A 12.5% (30:0.8) SDS-polyacrylamide gel was loaded with purified VLPs from a T5620-8 extract (lane a), a total protein extract from T5620-8 (lane b) and a total protein extract from T91 (lane c). The ~68 kD p1-IFN fusion protein is indicated.

because the proteins have common antigenic determinants (Figure 7.2).

The antiserum that was obtained was tested in a Western blotting experiment (Figure 7.5). Four identical nitrocellulose filters were prepared: each had two tracks of transfered total cell protein extracts, one from a T91 (vector) culture and one from a T91-1 (interferon expressing, Mellor et al., 1985d) culture. The filters were probed with either an anti Ty-VLP antiserum (Adams et al., 1987) or the anti Ty:IFN-VLP antiserum. The prebleeds were also assayed to ensure that the antisera hybridisations were, indeed, specific reactions. The anti Ty-VLP antiserum reacted with all the Ty related proteins in the T91 extract, the same proteins were detected in the T91-1 extract indicating that no interferon related proteins had reacted. The anti Ty:IFN-VLP antiserum reacted with the same Ty related proteins in the T91 extract, however a previously undetected protein of ~20 kD was identified in the T91-1 extract. This is the interferon that is present in T91-1 extracts and clearly shows that the Ty:IFN hybrid VLPs can elicit an antibody response to interferon. Therefore, these hybrid Ty-VLPs may be regarded as self-assembling epitope carriers. These observations have been described by Malim et al. (1987b).

The abundance and distinctly particulate nature of the hybrid Ty-VLPs make them an attractive system for antigen purification and presentation. Whether the particles remain intact, once injected into rabbits, such that a high, local concentration of antigen is maintained, or whether they are broken down into protein aggregates with a reduced mass, is

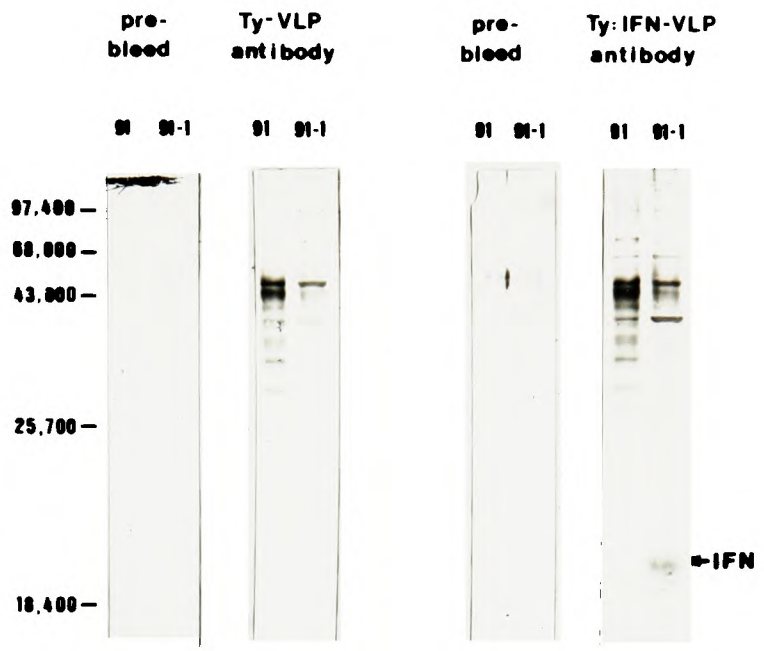


Figure 7.5. Western blot analysis of the total proteins from T91 and T91-1 extracts. The proteins were separated on a 12.5% (30:0.8) SDS-polyacrylamide gel and the filters were probed with an anti Ty-VLP antiserum and an anti Ty:IFN-VLP antiserum as well as the corresponding prebleeds. The specific hybridisation to the interferon present in the T91-1 extract is indicated.

currently unknown.

7.4 Hybrid VLPs Carrying a Viral Epitope

To demonstrate the applicability of the hybrid Ty-VLPs to a biological pathogen, a fragment from the haemagglutinin 1 (HA1) region of the HA gene of an H3N2 (H3 subtype) influenza virus was assayed for its ability to produce an antigenic fusion protein when inserted into pMA5620.

A 262 bp SpeI fragment that corresponds to the antigenic determinant of the HA1 polypeptide (codons 25-111, based on the data of Wilson et al., 1984) was isolated from the plasmid MX29 (refer to the legend to Figure 7.6), filled-in and ligated into the HincII site of pSP46. The relevant BamHI-BglII fragment was isolated from this and inserted into the BamHI site of pMA5620 so that the TYA and HA reading frames were in phase (Figure 7.6). The resulting plasmid was designated pMA5620-F and was transformed into MD40-4c.

Extracts of T5620-F were shown to contain a ~60 kD particulate protein (the predicted size of the TYA:HA fusion protein) by the methods described earlier (data not shown). The morphology was very similar to the T5620-8 "woolly" particles observed earlier (Figure 6.8). To test the efficacy of these hybrid Ty:HA-VLPs in eliciting an immune response to HA1, they were purified and rabbits were immunised with them. The antiserum that was obtained was tested in a Western blotting experiment (Figure 7.7). Duplicate filters with the blotted

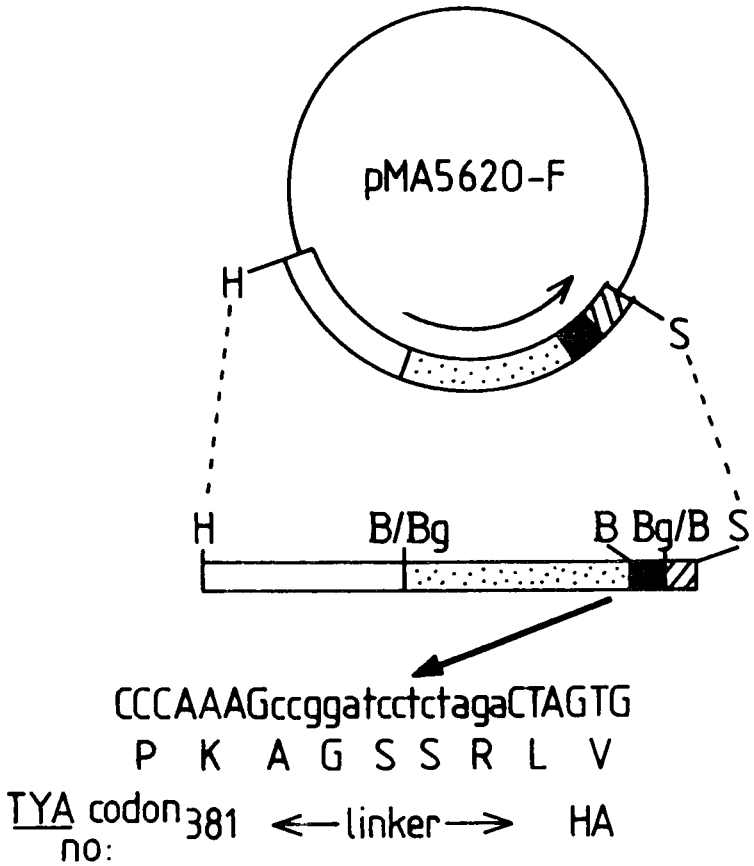


Figure 7.6. The structure of pMA5620-F and the nucleotide sequence and deduced amino acid sequence of the TYA:BamHI linker/pSP46 polylinker:HA junction. The symbols are described in the legend to Figure 7.1 except that the solid box corresponds to HA rather than IFN sequences. MX29 is a plasmid that contains HA1 and HA2 from the A/Memphis/102/72 (H3N2) strain of influenza virus; it was constructed by the GC tailing of a cDNA followed by insertion into the PstI site of pBR322.

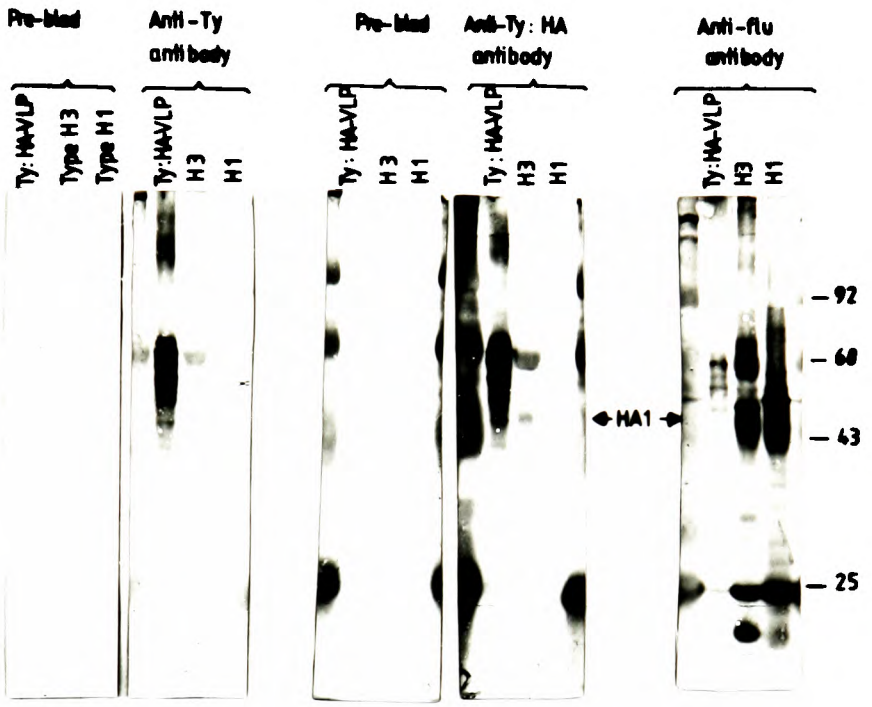


Figure 7.7. Western blot analysis of the proteins from purified Ty:HA-VLPs (from T5620-F extract) and purified influenza viruses (H1 and H3 subtypes). The proteins were separated on a 12.5% (30:0.8) SDS-polyacrylamide gel and the filters were probed with either an anti Ty-VLP antiserum, an anti Ty:HA-VLP antiserum, the corresponding prebleeds or an anti flu antibody (raised against the A/Scotland/840/74 x A/PR/8/34 (H3N2) strain of influenza virus, obtained from the World Influenza Centre at NIMR, Mill Hill, U.K.). The specific hybridisations to the HA1 polypeptide in the H3 subtype virus tracks are indicated.

proteins from purified Ty:HA-VLPs and purified influenza viruses (H1 and H3 subtypes) were prepared. These were hybridised with both an anti Ty-VLP antiserum and the anti Ty:HA-VLP antiserum, as well as the prebleeds and an anti flu antiserum. Hybridisation with both of the prebleeds produced a signal with a band of ~65 kD in the H3 subtype virus tracks, this is the result of non-specific association to a highly abundant nuclear protein. The anti Ty-VLP antiserum did not react with any additional proteins in either of the virus tracks. The anti Ty:HA-VLP antiserum did, however, react with a ~50 kD protein, but only in the H3 subtype virus track: this is the haemagglutinin 1 polypeptide. The lack of any reaction to a similarily sized protein in the H1 subtype virus track demonstrates the subtype specificity of the antiserum. The reaction with the anti flu antiserum was a positive control that hybridised to all the proteins in both of the lanes containing purified viruses.

Therefore, hybrid Ty-VLPs are able to function as self-assembling epitope carriers, when both a model antigen (α2-interferon) and a viral antigen (haemagglutinin 1) are fused to a truncated p1 protein.

7.5 Conclusions and Discussion

The ability of a truncated TYA gene, when fused to an ORF (or part of an ORF), to yield a fusion protein that assembles into hybrid Ty-VLPs has been demonstrated. These particles can be obtained in high yields (for example, 2 mg protein/10^{10} cells) using an easy purification procedure; furthermore, they are able

to elicit an immune response, in rabbits, to the epitopes of both the Ty and ORF encoded protein components. This has been shown with an α2-interferon cDNA and a part of the haemagglutinin 1 gene of a H3 subtype influenza virus.

The results presented in this chapter are of a preliminary nature. To date, the antisera that were produced have not been compared with antisera raised by alternative methods, for example, one against purified interferon itself. Also, to gain a better evaluation of the effectiveness of the antisera, they need to be tested in solution phase assays such as immunoprecipitations. This would determine whether or not the ORF related antigen was being presented to an immune system in its native conformation. Nevertheless, these results suggest that the insertion of any ORF into the particle expression vector, pMA5620, will lead to the production of hybrid particles that can elicit an immune response to the peptide for which the ORF codes: this is pictorially represented in Figure 7.8. Therefore, Ty-VLPs can function as self-assembling epitope carriers in a manner similar to the HBsAg fusions that were described earlier: consequently, they may have applications in industrial, medical and experimental fields such as vaccine development and monoclonal antibody screening.

Since the completion of this work, antisera against the env protein of HIV (Adams et al., submitted) and the E1 and E2 proteins of Bovine Papilloma Virus (BPV) (Harrison et al., unpublished data) have been raised using hybrid Ty-VLPs.

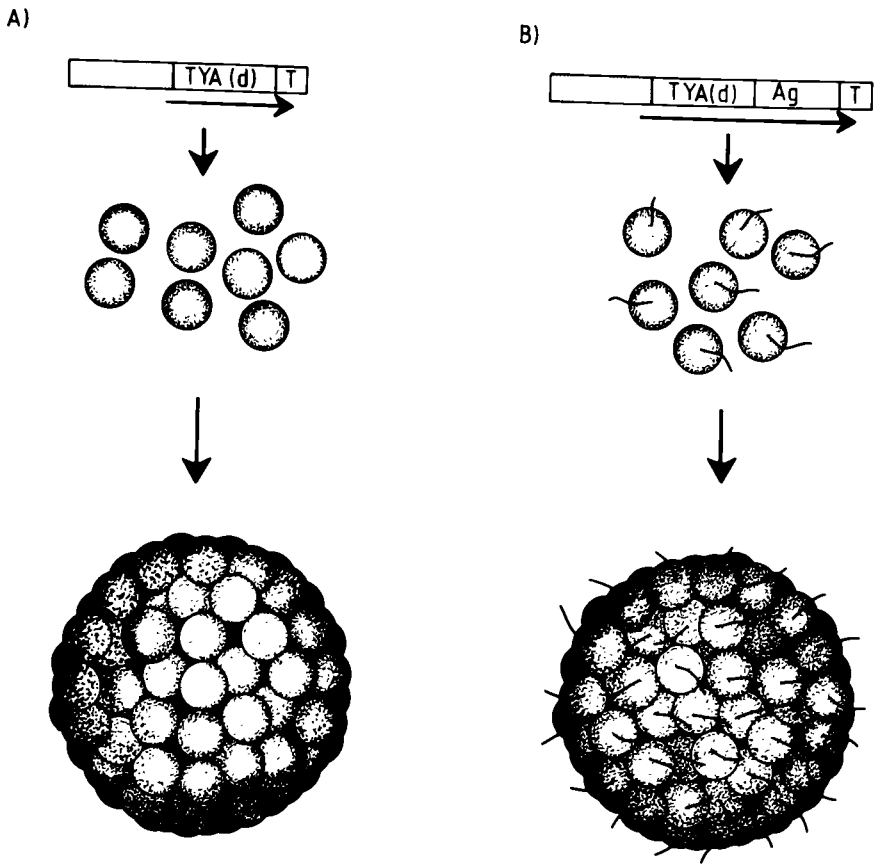


Figure 7.8. Pictorial representation of Ty-VLP (A) and hybrid Ty-VLP (B) formation. Shaded box, 5' promoter region of the yeast PGK gene; TYA (d), truncated TYA gene; T, 3' terminator region of the yeast PGK gene; Ag, any biological antigen (for example, α2-IFN or HA1).

CHAPTER 8. A SEARCH FOR THE cDNA INTERMEDIATES OF TY ELEMENT TRANSPOSITION

8.1 Introduction

It has been described in Chapters 1 and 3 how eukaryotic retrotransposons, such as Ty and copia, resemble integrated retroviral proviruses. These elements share similar genome organisations, consensus sequences, expression strategies and protein functions. For example, virus-like particles containing Ty and copia RNA, along with an associated reverse transcriptase, have been identified in both yeast cells and cultured Drosophila cells, respectively (Mellor et al., 1985c; Garfinkel et al., 1985; Shiba and Saigo, 1983). These observations have led to the prediction that the life cycles of retrotransposons resemble those of retroviruses (Figures 1.5 and 3.10) (Panganiban, 1985). Reverse transcription of the genomic retroviral RNA yields a linear dsDNA copy that has two LTRs and represents the whole genome but with two additional base pairs at each end (Figure 1.5). These two ends are blunt-end ligated to each other to produce a 2-LTR circle which can be integrated into the host genome by non-homologous recombination. This requires site-specific endonucleolytic cleavage within att (LTR-LTR junction) and the insertion site, these processes are mediated by the retrovirally encoded integrase (Panganiban and Temin, 1984; Duyk et al., 1985; Grandgenett et al., 1986). The participation of RNA intermediates in both Ty element transposition and retroviral transmission has been demonstrated (Boeke et al., 1985; Linial, 1987).

The extrachromosomal retroviral DNA of virally infected cells has been characterised. There are both linear and

supercoiled species (Shank et al., 1978). The supercoiled species comprise both one and two LTR circles, with considerable sequence variation observed at the LTR-LTR junction of the 2-LTR circles (Van Beveren et al., 1982). The 2-LTR circles are produced predominantly by reverse transcription and the subsequent ligation of the two termini (see above). The 1-LTR circles are thought to be the result of homologous recombination between the LTRs of either a 2-LTR circle or a genomic provirus, and are therefore not the direct products of reverse transcription.

The purification of the extrachromosomal DNA from cultured Drosophila cells has shown that both linear and circular species with homology to copia (Flavell and Ish-Horowicz, 1981 and 1983; Flavell, 1984), along with circular species with homology to the copia-like elements 297, mdg1, pDm39 and gypsy (Mossie et al., 1985) are present. The relative abundance of these various circles is not correlated to the number of genomic copies, or the poly A^+ RNA levels, of that repetitive element; however, there is typically one copia-circle in 1-50 cells. As with retroviral extrachromosomal DNA, the copia-circles have either one or two LTRs. The 1-LTR circles are likely to be the result of homologous recombination. The 2-LTR circles, which may have small inserts at the LTR-LTR junction, are probably the result of imprecise excision (in the case of those with inserts) or reverse transcription. The termini of some linear-copias have been sequenced and consequently shown to be exact copies of integrated elements (Flavell and Brierley, 1986). The self-ligation of these

linears would produce an LTR-LTR junction with no intervening nucleotides, this species has been proposed as an intermediate of copia transposition. This precise LTR-LTR junction is unlike that of the reverse transcriptase generated 2-LTR circle junctions of retroviruses, which have intervening nucleotides (Figure 1.5), and implies that the copia cDNA is either primed in a sightly different way or is modified prior to its self-ligation. It should be stressed, however, that the inclusion of an RNA intermediate in copia element transposition has yet to be proven.

It therefore seems likely that extrachromosomal DNA molecules with homology to Ty elements should exist in yeast cells. Some of these would be the transposition intermediates produced by the reverse transcription of the 5.7 kb Ty RNA that is present in Ty-VLPs (Figure 3.3). Attempts have been made by several investigators to identify such species by the preparation of extrachromosomal DNA followed by hybridisation analysis. However, such experiments have been unsuccessful to date (various personal communications).

Therefore a strategy that utilised a selection assay in E. coli rather than detection by hybridisation was adopted; it was hoped that this would enhance the likelihood of finding such DNA species. It was assumed that any intermediate DNA species would be circular at some stage, so that the inclusion of an E. coli origin of replication would render it capable of autonomous replication when isolated from yeast and introduced into E. coli.

The approach resembled the one that was used to study the packaging of Ty RNA into Ty-VLPs (Chapters 4 and 5). A "helper"

plasmid (pMA132-1Ø) was cotransformed into yeast with a second plasmid that was designed to produce a modified RNA. This RNA, inspite of being expression deficient, should be packaged into Ty-VLPs, reverse transcribed by the enzyme present in the "helper" Ty-VLPs and circularised to yield a cDNA capable of conferring antibiotic resistance when transformed into E. coli.

The E. coli origin of replication and selectable marker, the ampicillin resistance gene, were derived from pBR322. This module was fused to the 3' end of the 11 fragment so that the RNA would be packaged into Ty-VLPs (Section 4.3). In addition, it was present in both orientations so that the effect of any fortuitous yeast terminators in one orientation, that would produce a prematurely truncated RNA, would not prevent the desired RNA species being produced. This internal region, or "epsilon-equivalent" had the authentic 3' LTR of Tyl-15 at the 3' terminus, and either the Tyl-15 5' LTR or the PGK promoter (in the pMA91 configuration) at the 5' terminus. To maximise the chances of detection, these plasmids possessed restriction sites (HindIII and EcoRI) that would be absent from the proposed circular intermediates. Thus, the restriction digestion of the extrachromosomal DNA from these cotransformants would render the plasmids, but not the proposed intermediates, incapable of transforming E. coli as a result of their fragmented nature.

8.2 Plasmid Constructions

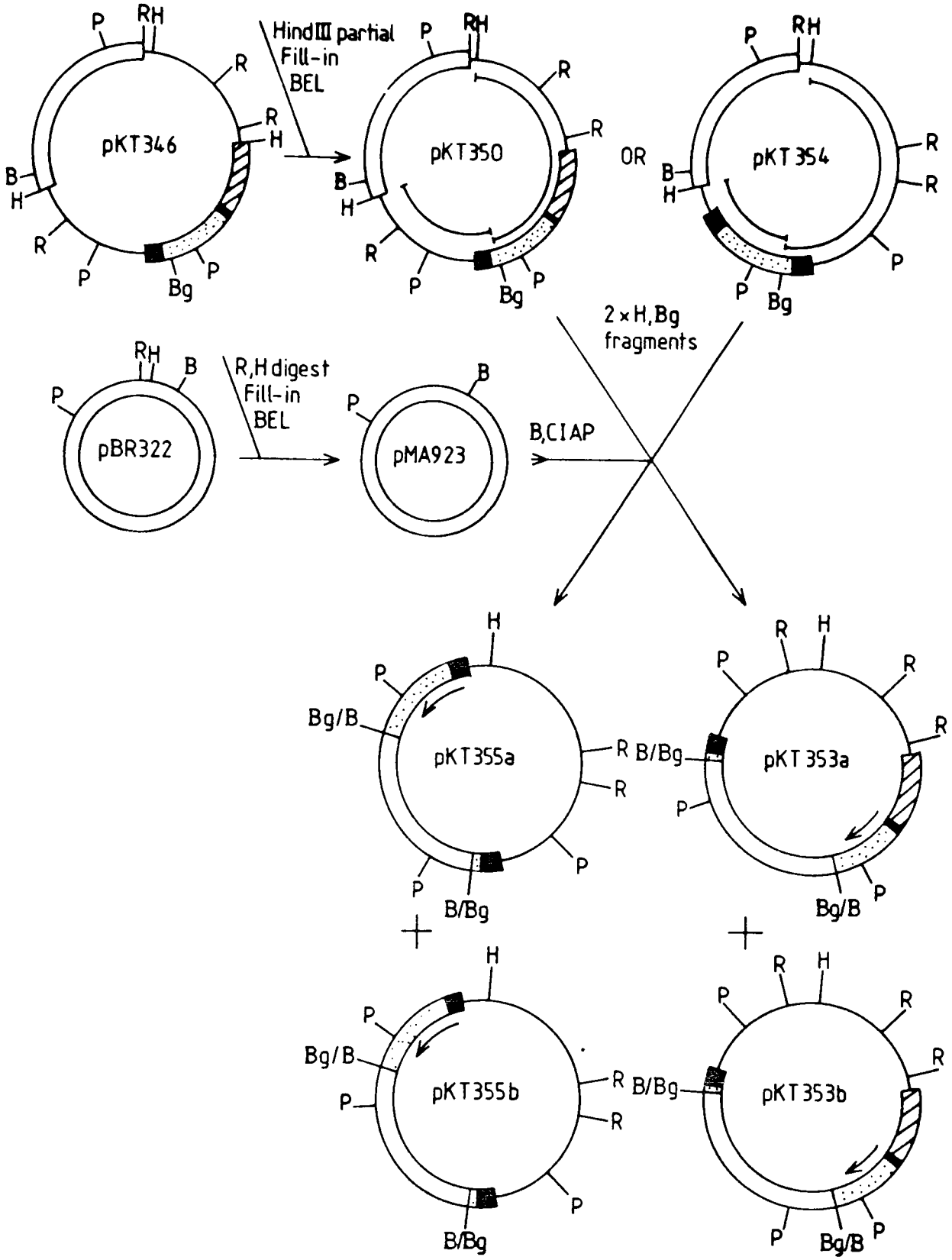
The scheme for constructing the PGK directed plasmids (pKT353a and pKT353b) is shown in entirety, whilst the construction of the 5' LTR directed plasmids (pKT355a and pKT355b) is somewhat abbreviated (Figure 8.1), by the omission of the manipulations that were required to produce pKT354.

The first step was to remove the HindIII and EcoRI sites from pBR322. This was achieved by double enzyme digestion followed by a fill-in reaction and blunt-end ligation. The resulting plasmid was designated pMA923, it has a unique BamHI site as well as an E. coli origin of replication and the Amp^R gene.

pKT346 was the starting molecule for the construction of pKT353a and pKT353b; it is the same as pKT346-8 (Figure 4.2) except that the IFN-8 fragment is absent, it therefore contains the PGK 5' promoter region and the Ty1-15 sequences that are 5' to the 5' BglII site and 3' to the 3' BglII site (including the 3' LTR). The HindIII site at the 5' end of the PGK promoter fragment was destroyed by a HindIII partial digestion, followed by the purification of the singly digested ~14 kb linear fragment, a fill-in reaction and its blunt-end ligation on itself. The resulting plasmid was designated pKT35Ø. The analogous plasmid for the construction of pKT355a and pKT355b was pKT354 (Figure 8.1); pKT354 has the same Ty derived sequences as pKT219 (Figure 4.2).

The final plasmids were constructed by the sticky end ligation of 3 fragments (Figure 8.1). pMA923 was linearised with

Figure 8.1. The construction of pKT353a, pKT353b, pKT355a and pKT355b. Thin lines, LEU2/2μ yeast selection and replication sequences and unspecified yeast sequences flanking Tyl-15 (similar to those of pKT227, Figure 4.6); open boxes, E. coli selection and replication sequences (pBR322 derived); hatched boxes; 5' promoter region from the yeast PGK gene; solid boxes; LTR derived sequences of Tyl-15; stippled boxes, sequences from the epsilon region of Tyl-15; R, EcoRI; H, HindIII; Bg, BglII; B, BamHI. The HindIII-BglII fragments used for the three fragment ligations are indicated, as are the directions of transcription of the Ty-pMA923 hybrid RNA transcripts of pKT353a, pKT353b, pKT355a and pKT355b.


pKT346
HindIII partial
Fill-in
BEL
pKT350
OR
pKT354
2 × H, Bg
fragments
pBR322
R,H digest
Fill-in
BEL
pMA923
B,CIAP
pKT355a
pKT353a
pKT355b
pKT353b
Bg/B
B/Bg

BamHI and treated with phosphatase to generate a BamHI fragment incapable of self-ligation: the two HindIII-BglII fragments from either pKT35Ø or pKT354 were purified and are indicated, the pBR322 sequences were therefore removed from between the yeast selection and replication module (LEU2/2μ) and the 3' LTR flanking sequences of these plasmids. The E. coli transformants were selected for their resistance to ampicillin (selects for the presence of pMA923) and their prototrophy for leucine (selects for the presence of the LEU2/2μ module, the LEU2 gene is functional in E. coli). This double selection minimised the amount of miniprep screening. Plasmids with the pMA923 sequences in both orientations were selected (pKT353a, pKT353b, pKT335a and pKT355b) and cotransformed into MD4Ø-4c with pMA132-1Ø (the "helper" plasmid).

The predicted 2-LTR circular transposition intermediate from T355a/132-1Ø is shown in Figure 8.2; it has no HindIII or EcoRI sites.

8.3 The Isolation and Restriction Analysis of Plasmids that Confer Ampicillin Resistance

To confirm that the four defective Ty elements constructed in Section 8.2 were capable of producing the corresponding full length fusion RNA species, a Northern blotting experiment was carried out. Selective cultures of all four cotransformants were grown at 22°C (the temperature optimum for the Ty-VLP associated reverse transcriptase, refer to Section 3.4), the RNA was

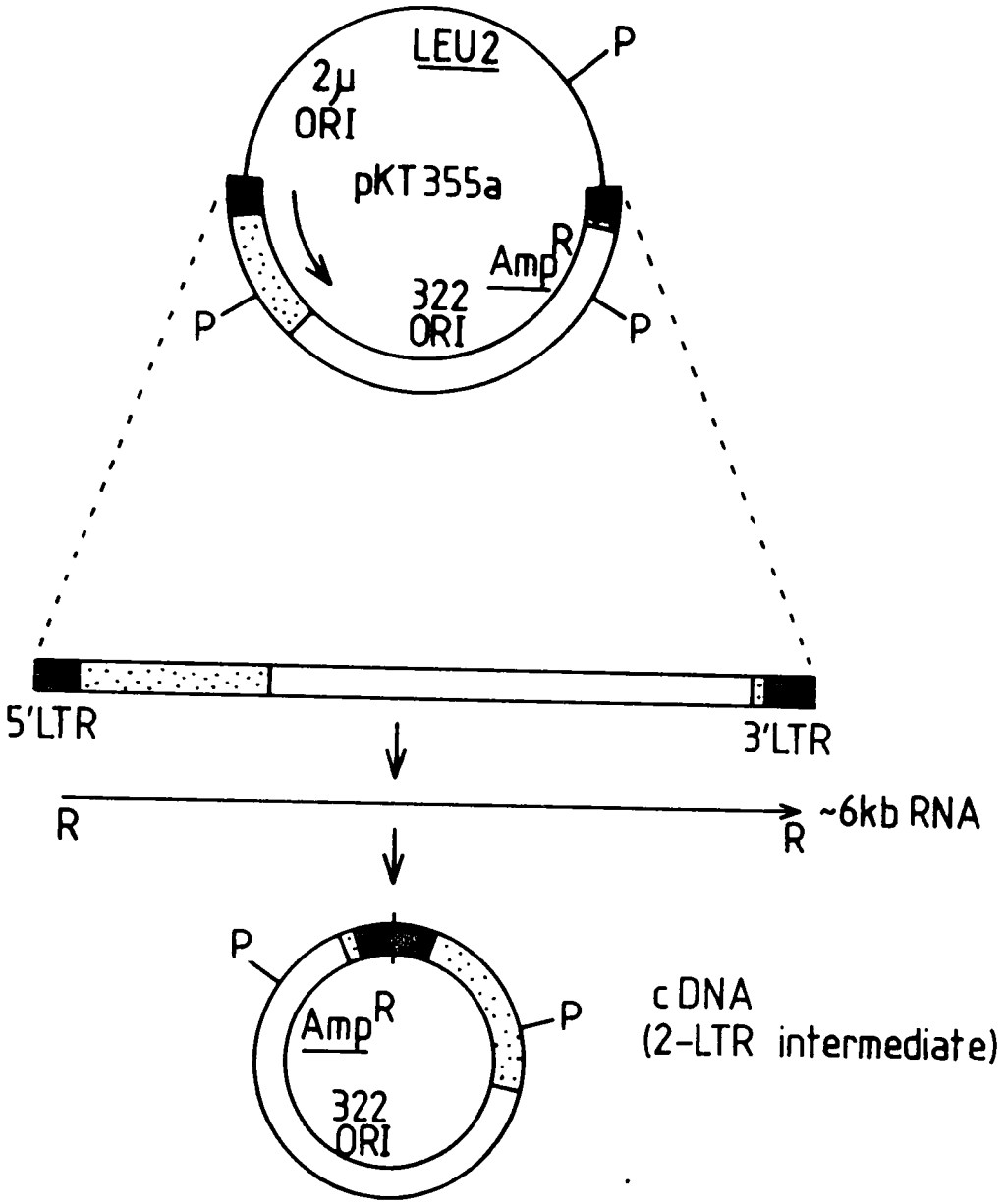


Figure 8.2. The structure of pKT355a and the proposed 2-LTR cDNA transposition intermediate that it should produce upon propagation in yeast cells that also contain pMA132-10. The symbols are described in the legend to Figure 8.1.

extracted and then analysed: the filter was probed with nick-translated pBR322 (Figure 8.3). The position of the ribosomal RNA occlusions allows the size approximation of the various RNA species. There is a ~6kb species, corresponding to the full length elements, in T353a/132-1Ø and T355a/132-1Ø. No such RNA species is present in T353b/132-1Ø or T355b/132-1Ø, indicating that a transcription terminator is present in the pBR322 derived sequences when the yeast promoter directed transcription proceeds from the $\underline{Amp}^R$ gene towards the $\underline{E.\ coli}$ origin of replication (refer to Figure 8.2). These two particular cotransformants would therefore be unable to produce the hypothesised transposition intermediates. The other bands are presumably various strand and antistrand transcripts from either one or both of the plasmids that have at least partial homology to the pBR322 sequences. For example, read through transcripts from $\underline{LEU2}/2\mu$ into pBR322 have been observed for similarily constructed plasmids (Kim, 1986).

In an attempt to isolate the proposed intermediates, 1 litre selective cultures of T353a/132-1Ø and T355a/132-1Ø were grown at 22°C and the supercoiled DNA was isolated using a CsCl gradient (Section 2.6.2). One quarter of each of these DNAs was digested with HindIII and EcoRI and used to transform $\underline{E.\ coli}$ (strain AKEC28) to ampicillin resistance. One transformant from the T353a/132-1Ø DNA and three from the T355a/132-1Ø DNA were obtained. Miniprep DNA was isolated from each of these $\underline{E.\ coli}$ transformants. These were digested with HindIII and EcoRI and, as expected, retained their supercoiled structure (data not shown). They were then digested with either PstI or XhoI (Figure

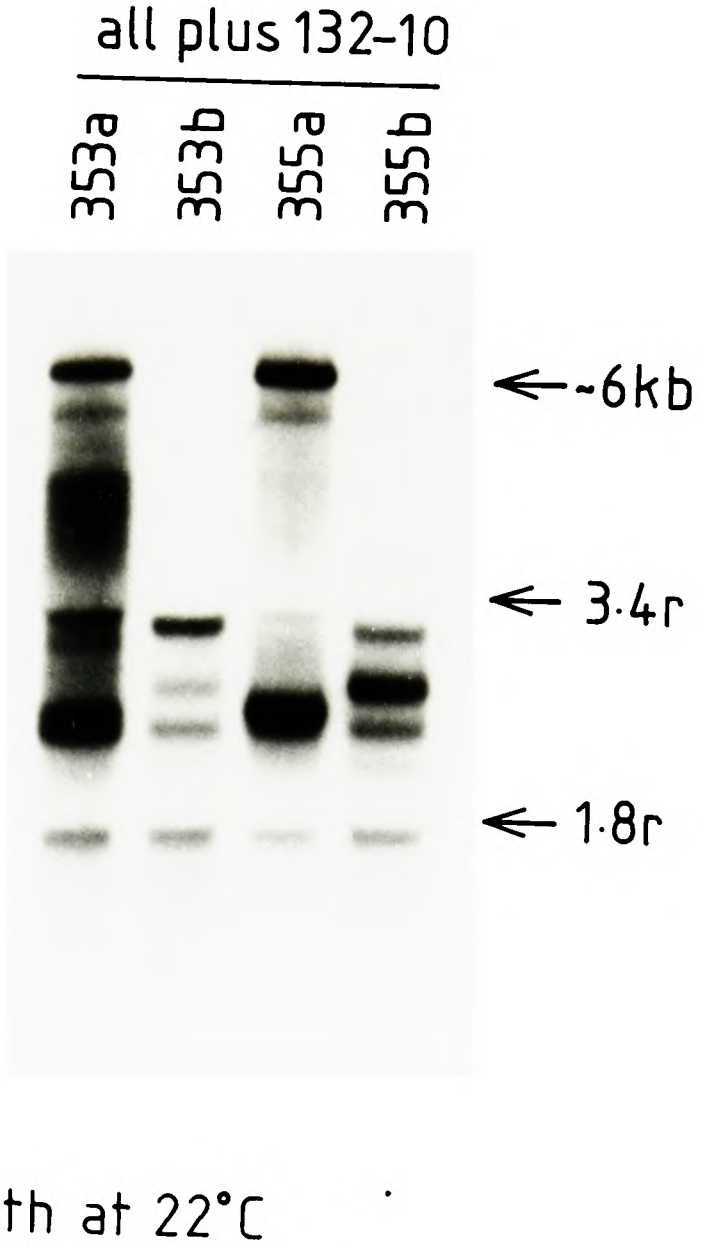


Figure 8.3. Northern blot analysis of the total cellular RNAs of T353a/132-1Ø, T353b/132-1Ø, T355a/132-1Ø and T355b/132-1Ø grown at 22°C. The hybridisation probe was pBR322. The positions of the ~6 kb RNA and the 3.4 kb and 1.8 kb ribosomal RNA occlusions are indicated.

8.4). PstI, as expected, restricted each DNA twice, however the smaller band was only ~2.2 kb rather than the ~2.55 kb that was expected if the DNA had the structure of the proposed 2-LTR transposition intermediate (Figure 8.4, lanes a, b, c and d). The T353a/132-1Ø derived DNA, as expected, had only one XhoI site (Figure 8.4, lane e); pKT353a only has the XhoI site of the 3' LTR since the site of the 5' LTR was modified for its fusion to the PGK promoter. The T355a/132-1Ø derived DNAs also had single XhoI sites (Figure 8.4, lanes f, g and h). pKT355a, however, has two intact LTRs and therefore has two XhoI sites, hence the predicted 2-LTR intermediate would also have two XhoI sites. Consequently, XhoI digestion would generate a fragment of 34Ø bp (the size of a single LTR); however, there is no evidence of such a DNA fragment.

It can be concluded that the supercoiled DNA species that were obtained did not have the structure of the predicted 2-LTR transposition intermediates. A likely explanation as to their origin, that would fit the restriction digestion data, is that these DNAs were the products of homologous recombination events across the repeated LTR sequences of pKT353a and pKT355a. Indeed, the size of the smaller, ~2.2 kb PstI fragment (Figure 8.4, lanes a, b, c and d) correlates with this conclusion. These species are analogous to the 1-LTR copia circles that were described earlier. Nevertheless, other mechanisms like anomalous tRNA binding, and therefore reverse transcriptase priming, can not be ruled out as possible explanations.

Some HB1Ø1 competent cells were purchased from BRL (Section

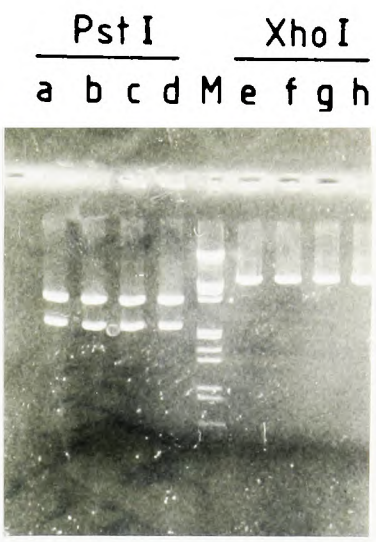


Figure 8.4. A 1% Tris-borate agarose minigel showing miniprep DNA of the T353a/132-1Ø DNA derived AKEC28 E. coli transformant digested with PstI (lane a) and XhoI (lane e) and the T355a/132-1Ø DNA derived AKEC28 E. coli transformants digested with PstI (lanes b, c and d) and XhoI (lanes f, g and h). The markers (M) were described in the legend to Figure 4.7.

2.11.2) in an attempt to maximise the transformation efficiency of the HindIII, EcoRI treated DNAs. All the transformants that were obtained contained DNA with the same restriction patterns as the DNAs obtained from the AKEC28 transformations (data not shown).

8.4 Discussion

These attempts to isolate the DNA intermediates of Ty element transposition were therefore unsuccessful. There are a number of possible explanations for this. Firstly, the experimental approach may still be insufficiently sensitive for the detection of a 2-LTR circular DNA species, for example, such species may have an extremely short life span. Secondly, the proposed 2-LTR circle may not actually exist: transposition may occur via the linear 2-LTR species. Indeed, recent evidence from other laboratories suggests that the transmission of retroviral sequences may not proceed via the 2-LTR circle (A. Kingsman, personal communication). Thirdly, it is possible that the yeast strain MD4Ø-4c possesses a genetic background that prevents Ty element transposition, either by disrupting the formation of, or inhibiting the chromosomal insertion of, the 2-LTR circles. For instance, a deficiency in the ligase activity that fuses the two free ends of a 2-LTR linear cDNA together would disrupt the formation of the 2-LTR circles. Furthermore, the frequency of Ty element transposition has, in fact, been shown to vary considerably within different strains of S. cerevisiae (S. Kingsman, personal communication).

The supercoiled species that were isolated were probably the result of LTR-LTR homologous recombination. This is the mechanism by which the hundred or so solo-deltas of the yeast genome were thought to have arisen (Cameron et al., 1979). The question still remains as to how the DNA transposition intermediates could be isolated. The experiments that are most likely to be successful should probably utilise yeast strains with proven high frequencies of Ty element transpostion, followed by extrachromosomal DNA isolation and either optimally sensitive nucleic acid hybridisation technology or some kind of selection procedure similar to the one described here. The latter approach might be improved by preventing the translation of the hybrid RNA (provided that its stability was maintained), thus all the Ty-VLPs would be supplied by the "helper" and would therefore contain reverse transcriptase. Another possible modification would be the insertion of a yeast ARS (autonomously replicating sequence) between the LTRs of the "assay" plasmids. Any intermediate species that is formed might then have a more sustained life span within a yeast cell.

CHAPTER 9. THE *TYB* GENE OF THE CLASS II ELEMENT TY1-17 IS EXPRESSED AS A *TYA*:*TYB* FUSION PROTEIN

9.1 Introduction

It was described in Section 1.4.5 how the TYB gene of the class I Ty element Ty1-15 is expressed as a TYA:TYB fusion protein of ~190 kD (Mellor et al., 1985a; Clare and Farabaugh, 1985; Wilson, et al., 1986). The TYB gene overlaps TYA by 38 nucleotides and is in the +1 reading phase with respect to TYA. It has been shown that an RNA splicing event is not responsible for the synthesis of the fusion protein (Wilson et al., 1986), therefore the most likely explanation is one of ribosomal frameshifting. This would bring the two out of phase reading frames into phase, thus avoiding translation termination by the TYA stop codon. This event occurs in approximately 10% of all the TYA translation events. Ribosomal frameshifting is also responsible for the synthesis of the $Pr180^{gag-pol}$ fusion protein of RSV (Jacks and Varmus, 1985), however, the two mechanisms are not directly analogous since the pol gene of RSV is in the -1 reading phase with respect to gag.

A region of 31 nucleotides within the 38 bp TYA/TYB overlap region of Ty1-15 is still capable of directing normal levels of frameshifting (Wilson et al., 1986). The genetic organisation of the class II Ty element Ty1-17 is very similar. TYB starts with a GCG (alanine) codon (Fulton et al., 1985) and overlaps TYA by 44 bp so that it is in the +1 reading phase with respect to TYA. The overlap regions of the two elements contain a completely conserved 11 bp region, but otherwise share very little homology; the deletion of 5 of the 11 conserved nucleotides from Ty1-15 results in the complete loss of frameshifting (Wilson et al.,

1986). Therefore the ability of the class II element Ty1-17 to produce an analogous TYA:TYB fusion protein was investigated. α2-interferon cDNA fragments were inserted into all three reading phases downstream from the TYA stop codon, so that a shift in codon recognition into the TYB reading phase could be detected by the production of a TYA:TYB-IFN fusion protein in a Western blotting experiment. Mellor et al. (1985a) have used a similar strategy to analyse the expression of the TYB gene of Ty1-15.

9.2 Plasmid Constructions

The scheme for the plasmid constructions is outlined in Figure 9.1. The starting molecule was pKT47. This is a high copy number plasmid that replicates in both E. coli and yeast, it contains the 5' half of Ty1-17 (from the 5' LTR to the HindIII site in the middle of TYB). To create the deletion endpoints necessary for the insertion of IFN cDNAs into all three reading phases, pKT47 was digested with BamHI (the loss of the small BamHI fragment was inconsequential) and treated with Bal31 exonuclease. The digested DNAs were filled-in, blunt-end ligated in the presence of excess BglII linkers (CAAAAGATCTTTTG; Celltech) and transformed into E. coli. Miniprep DNAs isolated from the transformants were analysed by digestion with HaeIII and BglII followed electrophoresis on polyacrylamide gels (data not shown). The deletion endpoints that were in the required region were determined by DNA sequencing. To do this HaeIII,BglII digested miniprep DNA was ligated into M13mp19 DNA digested with

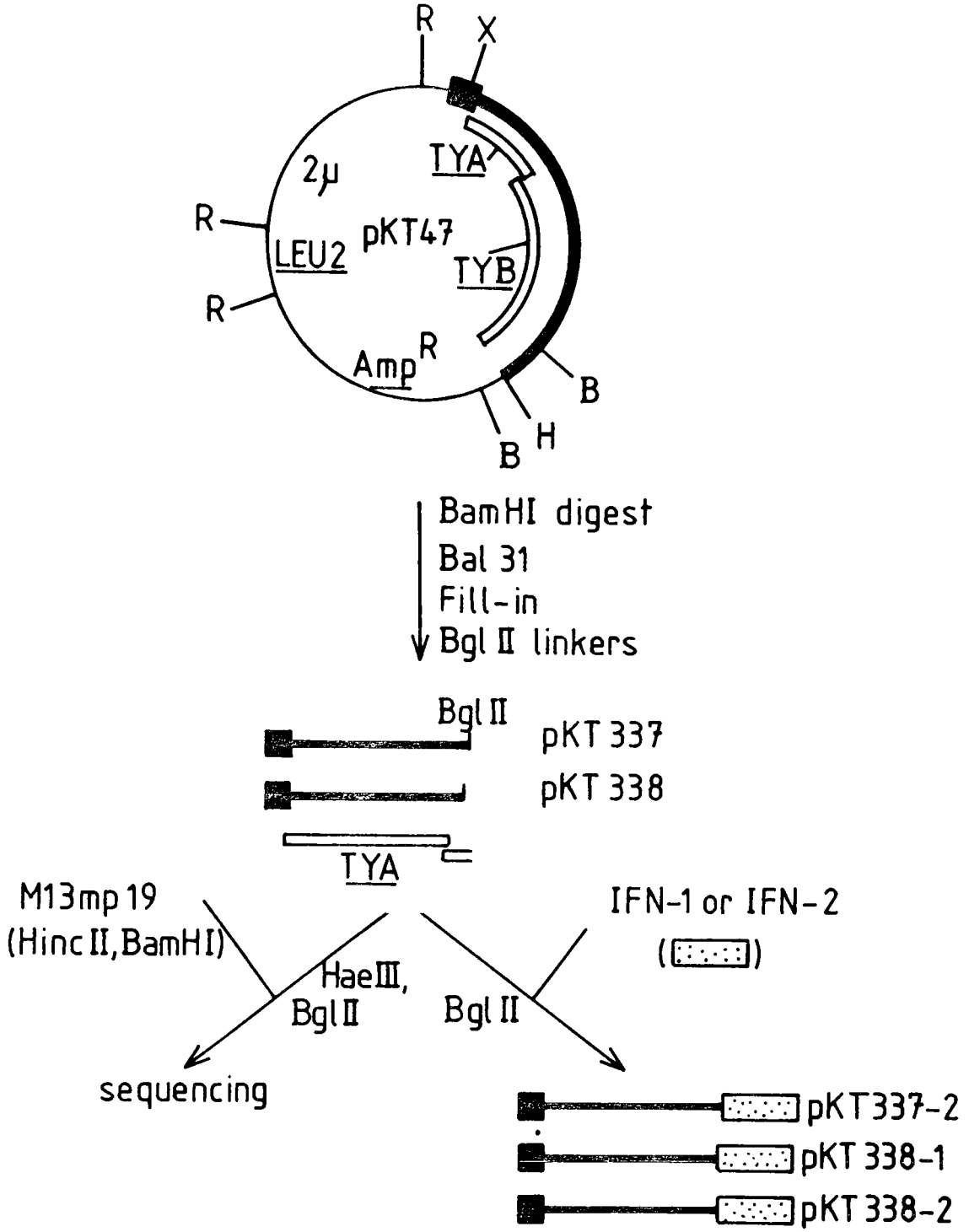


Figure 9.1. The construction of pKT337-2, pKT338-1 and pKT338-2. Thin line, yeast/*E. coli* selection and replication sequences and unspecified sequences flanking Ty-17; solid boxes, 5' LTR of Tyl-17; thick lines, epsilon region of Tyl-17; open boxes, Tyl-17 ORFs (*TYA* and 5' half of *TYB*); stippled boxes, IFN-1 and IFN-2 fragments; R, EcoRI; X, XhoI; B, BamHI; H, HindIII.

HincII and BamHI. Two endpoints that were 79 and 75 nucleotides beyond the TYA stop codon were selected and the plasmids designated pKT337 and pKT338, respectively. The IFN-1 and IFN-2 fragments (Figure 4.3) were inserted into the unique BglII sites of these plasmids to create pKT337-2, pKT338-1 and pKT338-2, such that the IFN cDNA reading frame was in all three reading phases. These three plasmids were transformed into MD40-4c.

9.3 Western Blot Analysis of the TYB-IFN Fusions

To ensure that the plasmids had maintained the correct TYB:BglII linker:IFN junction, once present in yeast, the plasmids were rescued (Section 2.6.1) and transformed in E. coli. Miniprep DNAs were isolated, digested with BglII and TaqI and run on a polyacrylamide gel (Figure 9.2A). The ~280 bp BglII-TaqI fragments that encompassed the junction were purified (Figure 9.2B) and ligated into M13mp18 DNA digested with BamHI and AccI, for DNA sequencing (Figure 9.3). The nucleotide sequences of the junctions from all three plasmids are shown in Figure 9.4, as is the predicted amino acid sequence when the TYB and IFN reading frames are in phase with each other (pKT337-2).

To assay for the presence of TYA:TYB-IFN fusion proteins in the above transformants, a Western blotting experiment was carried out. The proteins from extracts of T337-2, T338-1, T338-2 and T361-1 were separated on an SDS-polyacrylamide gel and transferred to nitrocellulose. The T361-1 strain contains a plasmid that produces high levels of an ~60 kD PGK-IFN fusion protein, this acts as a molecular weight marker and a positive

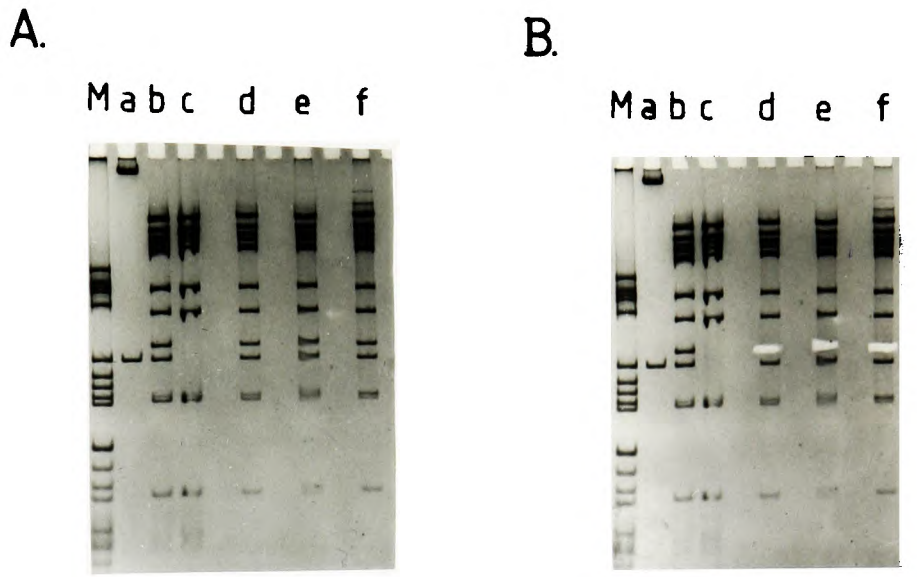


Figure 9.2. A, 7.5% (29:1) polyacrylamide gel showing BglII, TaqI digested miniprep DNA isolated from E. coli transformed with rescued plasmid DNA from T337-2 (lane d), T338-1 (lane e) and T338-2 (lane f). The markers (M) are pBR322 DNA digested with HaeIII; the band sizes (in bp) are 587, 540, 504, 458, 434, 267, 234, 213, 192, 184, 124 and 123 (appear as a doublet), 104, 89, 80, 64, 57 and 51. The remaining lanes contain pKT337-2 digested with BglII (lane a), BglII, TaqI (lane b) and TaqI (lane c). B, the isolation of the ~280 bp BglII-TaqI fragments from lanes d, e and f for ligation into M13mp18 DNA digested with BamHI and AccI.

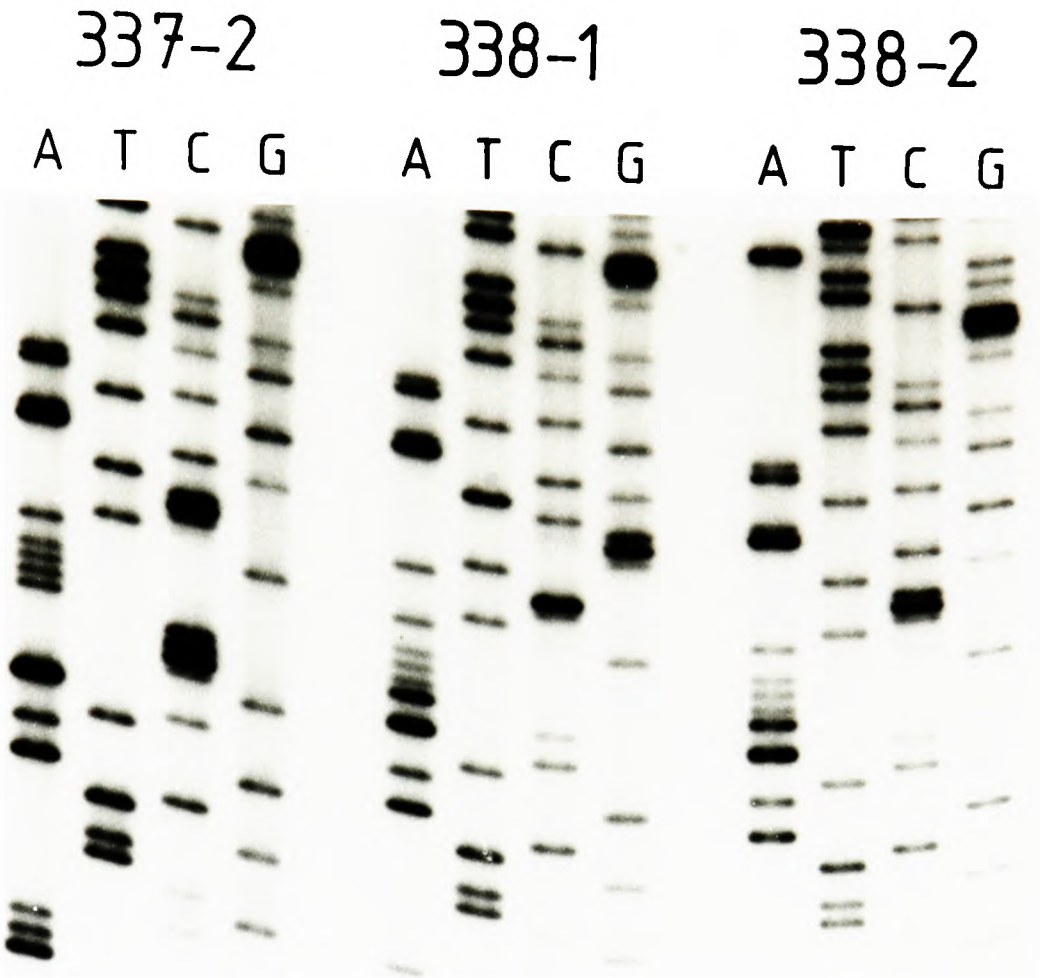


Figure 9.3. "Dideoxy" sequencing analysis of the TYB:BglII/BamHI linker:IFN junctions of pKT337-2, pKT338-1 and pKT338-2.

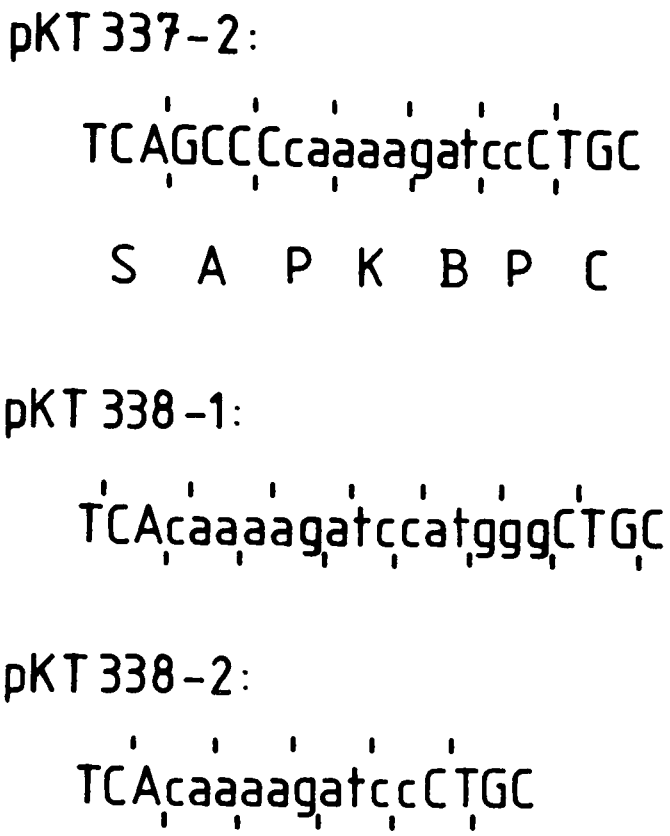


Figure 9.4. The nucleotide sequences of the TYB:BglII/BamHI linker:IFN junctions of pKT337-2 (also shows the deduced amino acid sequence), pKT338-1 and pKT338-2, as seen in Figure 9.3. Upper case nucleotides are derived from TYB and IFN; lower case nucleotides are derived from the linker sequences; lines above the nucleotides indicate the IFN reading phase; lines below the nucleotides indicate the TYB reading phase. The two phases only coincide in pKT337-2.

control for hybridisation. The filter was hybridised with an ^{125}I labelled monoclonal anti interferon antibody and exposed to X-ray film (Figure 9.5). The T361-1 extract (lane a) contains the ~60 kD PGK-IFN fusion protein as well as some smaller breakdown products. The T337-2 extract (lane b) contains an ~70 kD protein that reacts with the anti interferon antibody: no such protein is present in the T338-1 (lane c) or T338-2 (lane d) extracts. For a protein of ~70 kD to be synthesised by T337-2, TYA (50 kD) must be expressed (there is no significant coding capacity in reading phases 2 or 3) and fused to IFN (20 kD) via the 5' end of TYB, with which the IFN is in phase. Therefore a shift in codon recognition from the TYA gene to the TYB gene results in the TYB gene of Ty1-17 being expressed as a TYA:TYB fusion protein.

9.4 Discussion

The data in this chapter demonstrates that the class II Ty element Ty1-17 expresses its TYB gene as a TYA:TYB fusion protein. Class I elements express their TYB genes in the same way, therefore the TYB expression strategy of both classes of element is conserved. A region of 31 nucleotides in the TYA/TYB overlap region of Ty1-15 is sufficient for frameshifting (Wilson et al., 1986); this region shares little homology with the overlap of Ty1-17, though there is a perfectly conserved region of 11 bp. This small region is essential for the production of the Ty1-15 fusion protein, by a mechanism that is thought to

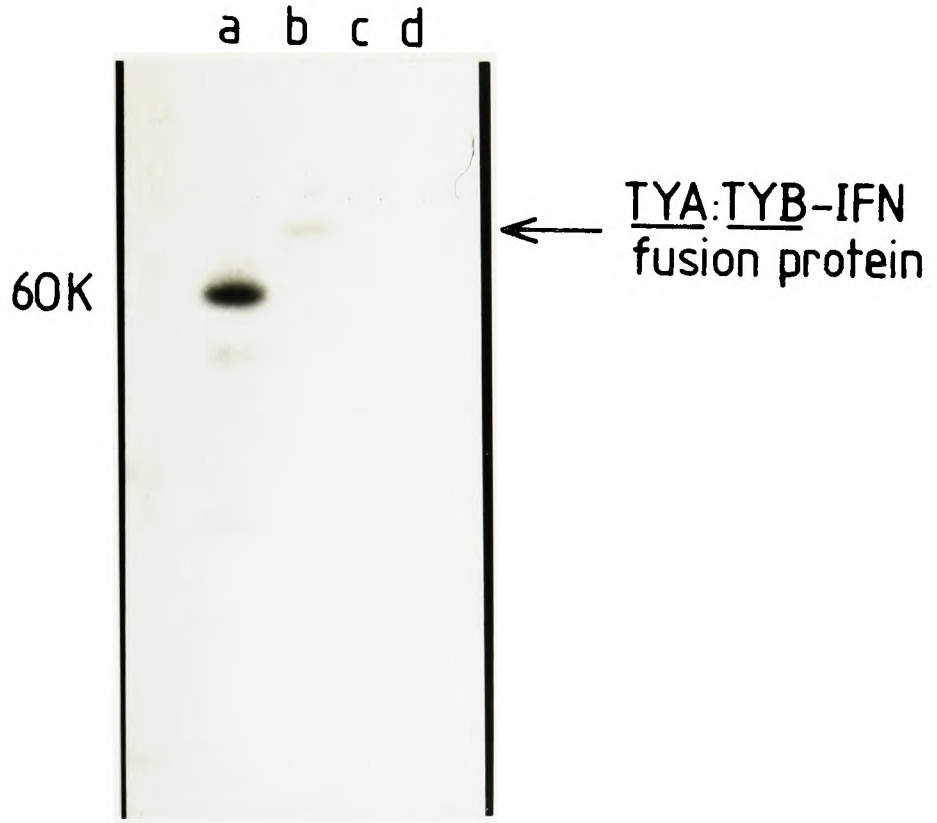


Figure 9.5. Western blot analysis of the total protein extracts from T361-1 (lane a), T337-2 (lane b), T338-1 (lane c) and T338-2 (lane d). The proteins were separated on a 12.5% (30:0.8) SDS-polyacrylamide gel and the filter was probed with an ^{125}I labelled monoclonal anti interferon antibody. The ~70 kD TYA:TYB-IFN fusion protein is indicated.

involve ribosomal frameshifting. The proposal of more detailed mechanistic models awaits the amino acid sequencing of the TYA:TYB junction of the fusion protein.

CHAPTER 10. DISCUSSION

10.1 General Discussion and Evolutionary Considerations

The similarities in genome organisation, expression strategy and protein function between the Ty elements of yeast and other eukaryotic retrotransposons such as copia, IAP genomes and TED, as well as the integrated proviruses of mammalian and avian cells, were described in Chapter 1. The natural progression from these observations was the prediction that Ty encoded proteins, like those encoded by RNAs homologous to copia and IAP genomes, would assemble into virus-like particles (VLPs) (Shiba and Saigo, 1983; Paterson et al., 1978). Electron micrographs of yeast cells overexpressing the whole transcriptional unit of the class I element Ty1-15 (T91-10 cells) identified an abundance of Ty-VLPs, with an average diameter of ~60 nm, within their cytoplasms (Figure 3.1). The subsequent purification and analysis of these particles demonstrated the association of both the 5.7 kb Ty RNA (genomic RNA) and a reverse transcriptase activity (Figure 3.3), whilst their major protein component was the Ty encoded, p2 (48 kD) (Figures 3.5 and 3.7).

Further experiments with Ty1-15 by Adams et al. (1987) showed that p2 and the less abundant Ty-VLP proteins p4, p5 and p6 are all proteolytic cleavage products of the primary translation product of TYA, p1 (50 kD). Adams et al. (1987) also proved that the proteolytic processing is mediated by a Ty encoded protease that is situated towards the 5' end of TYB. This processing pathway, together with the previous knowledge of Ty element expression strategy (Figure 1.7) is illustrated in Figure 10.1. This scheme shares many of its features with the

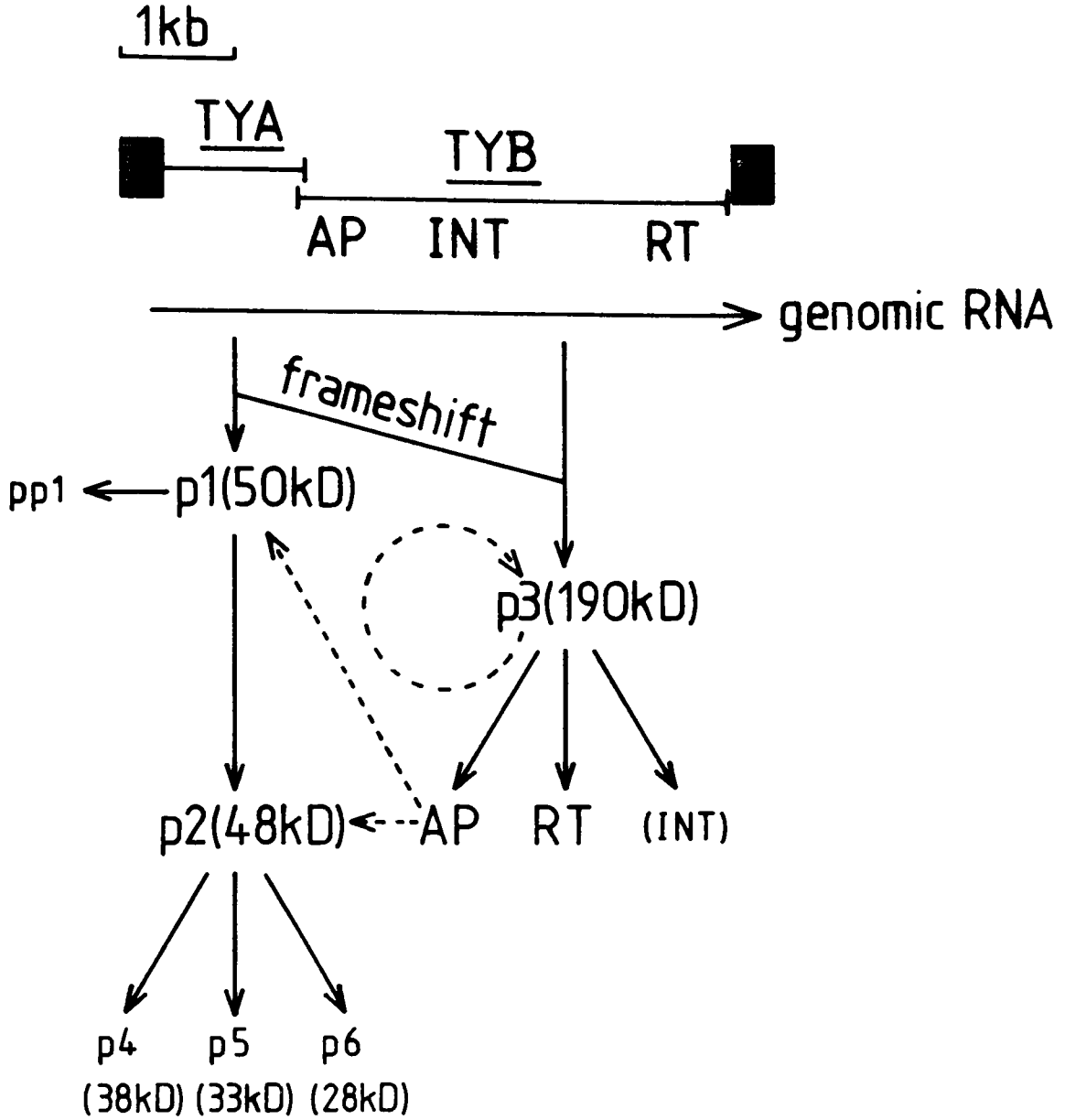


Figure 10.1. Genome organisation and expression strategy of a Ty element (1987). Solid boxes, LTRs; thin lines, ORFs; AP, INT and RT, homologies to the consensus sequences of retroviral acid proteases, integrases and reverse transcriptases; p, protein; pp, phosphoprotein; broken arrows, target sites for Ty encoded acid protease.

expression strategy of Avian Leukosis Virus (Figure 1.3), including the utilisation of polyprotein precursors, expressed both simply and as fusion proteins, and their extensive processing (by both proteolysis and phosphorylation). There are also similarities in the function of the mature proteins. These include being the structural components of viral particles, a reverse transcriptase and an acid protease. No identification of a Ty encoded integrase, as suggested by the presence of the consensus sequence within TYB (Mount and Rubin, 1985), has been made as yet.

The ability of p1 to assemble into particles (Figure 3.8) with a different morphology to T91-1Ø Ty-VLPs was also demonstrated (Figure 3.9; Adams et al., 1987). These observations, together with the knowledge that Ty elements transpose via RNA (Boeke et al., 1985), have led to the proposed Ty element life cycle that is outlined in Figure 3.1Ø. This predicts a requirement for the Ty-VLPs as the site of the reverse trancription that produces a DNA copy from the 5.7 kb Ty RNA template. The cDNA is predicted to enter the nucleus and integrate, more or less randomly, into the genome. An attempt was made to isolate the cDNA intermediates of transposition by producing modified Ty RNAs that should have been reverse transcribed and self-ligated into plasmids capable of propagation in E. coli (Chapter 8). This was unsuccessful for a number of possible reasons. The formation of the predicted 2-LTR cDNA may not be required for transposition, alternatively it could have a very short half life or MD4Ø-4c might inhibit its formation. The

apparent absence of Ty-VLPs from the nucleus does not constitute an argument against a role for Ty-VLPs in transposition since the DNA copy of retroviral RNA templates is transported to the nucleus, independently of virus particles, prior to chromosomal integration (Figure 1.5).

It was shown in Figure 6.2 that a p1 protein that had been trunated by 60 amino acids at its carboxyl terminus is still able to assemble into Ty-VLPs. The removal of a further 95 amino acids was sufficient to abolish particle formation (Figure 6.2). The fusion of either IFN or HA1 (H3 subtype virus) ORFs to the 3' end of the TYA gene truncated by 60 codons did not prevent the formation of particles (Figure 6.4 and data not shown). These hybrid Ty-VLPs elicited antibody responses in rabbits to both the Ty and ORF encoded epitopes (Figures 7.5 and 7.7), thus establishing them as self-assembling epitope carriers, the applications of which were discussed in Section 7.5.

A series of experiments were undertaken to identify the sequences within RNA homologous to Ty1-15 that are responsible for its incorporation (or packaging) into Ty-VLPs. An assay system was devised that relied upon an abundant PGK-Ty-IFN hybrid RNA being packaged, along with full length 5.7 kb Ty RNA, in the presence of Ty-VLPs that were provided by an efficiently expressed "helper" Ty element (Chapter 5). The Ty-VLPs were isolated and their RNA contents were analysed by Northern blotting; the amounts of the 5.7 kb Ty RNA and hybrid RNA, in the context of their cellular levels, were quantitated and compared to give a relative value for the packaging efficiency of the hybrid RNA, known as the Packaging Coefficient (Section 4.4).

The results obtained with a variety of hybrid RNAs indicated that Ty RNA packaging is a specific event but that the sequences responsible are dispersed throughout the genomic RNA (Figure 5.2). The efficient packaging of the PT357-1 "assay" RNA (Figure 4.11) was in contrast to the PT924-453 result (Figure 5.2) and may reflect the formation of a complex secondary RNA structure that is efficiently packaged in PT357-1 (Sections 4.6 and 6.5). The packaging signals within RNA homologous to the 11 fragment were concluded to be within the first 236 nucleotides (corresponds to the region 5' to the PvuII site of TYA) and downstream from nucleotide 737 (corresponds to the pKT359 deletion endpoint, Figure 4.8).

These findings were rather different from retroviral RNA packaging signals which are predominantly in the 5' non-coding and gag coding regions of the genomic RNAs. The env-mRNA and other subgenomic RNA species are not packaged into viral particles because at least some of the packaging signals have been removed by splicing. These RNAs are therefore not transmitted to the genomes of other cells by retroviral infection, reverse transcription and integration (Section 4.1). A plausible reason for the rather less localised nature of the Ty RNA packaging signals is suggested by the subgenomic Ty RNA species being the result of alternative transcription initiation and termination events (rather than RNA splicing) so that they have different 5' and 3' termini from the full length transcript (Section 1.4.3); the inclusion of the smaller RNA species into Ty-VLPs (unlike the env-mRNA into retroviral particles) can not

lead to their reverse transcription and chromosomal integration because of the absence of crucial sequences for these processes, such as the primer binding sites for the initiation of reverse transcription. Consequently, the selection pressure to conserve localised packaging signals is less stringent in Ty elements than retroviruses, providing that their specificity is maintained. The inability of RNAs with no homology to Ty elements to be packaged was demonstrated (Figure 4.5).

The efficient packaging of RNA with the authentic Ty termini (Figure 4.5) implies that the 5.7 kb Ty RNA, produced by the 30-35 endogenous Ty elements, might be packaged into Ty-VLPs as soon as a small amount of TYA translation is completed. Such an event would segregate the RNA from the ribosomes and therefore prevent translation. This could be a contributory factor towards the inefficient translation of Ty RNA; for example, a T91 cell has one twentieth of the Ty RNA of a T91-10 cell but its cytoplasm contains maybe only a thousandth of the number of Ty-VLPs (compare Figure 3.2 to Figure 3.1). This observation may support the earlier suggestion that the processes of translation and particle assembly might be coupled (Section 4.1). The relatively low levels of Ty proteins (for example, reverse transcriptase) could contribute towards the low frequencies of transposition that are observed in most yeast strains.

A study on the expression of the TYB gene of the class II element Ty1-17 showed that it is expressed as a TYA:TYB fusion protein (Figure 9.5), rather like the TYB genes of the class I elements (Wilson et al., 1986). This finding supports the prediction that both class I and class II elements share an

evolutionary ancestor, a fact that is borne out further by the assembly, into Ty-VLPs, of the Ty1-17 proteins. The conserved function of TYA encoded proteins thatwas proposed in Section 1.4.1 would therefore be particle formation. Why there are six times as many class I elements as class II elements may relfect the more successful propagation of class I elements or the creation of the "first" class II element at a time when there were already six other class I elements.

Although the existence and composition of copia-VLPs and IAPs have been recognised for some time, the processes involved in the expression of the constituent proteins and their subsequent assembly into VLPs, along with RNA, are poorly understood in comparison to the information regarding Ty-VLPs. For example, the mechanisms by which the reverse transcriptases that are associated with copia-VLPs and IAPs are produced, from an ORF that also codes for structural proteins and an ORF that is 3' to several stop codons (in the case of IAP-H18, at least), are unknown. Furthermore, unlike Ty element transposition (Boeke et al., 1985), the utilisation of an RNA intermedate in copia element and IAP genome transposition has yet to be demonstrated.

The unifying feature of all the retrotransposon VLPs examined so far is their inability to infect other cells. The restriction to intracellular targets is a likely consequence of the particles having no "env" protein. The env encoded surface glycoproteins of retroviruses confer both stability and host range specificity to the virus particles. For example, the glycoproteins of HIV restrict viral infection to cells of the

immune system and brain that express the CD4 molecule on their surface (Maddon et al., 1986). Thus, retroviruses can infect any cell that has the cognate cell-surface receptors (intergenomic transmission), whereas retrotransposons can only spread amongst the genome from which they originated (intragenomic transmission). An exception to this rule may be the Drosophila copia-like retrotransposon 17.6. This has both an "env" equivalent ORF and the characteristic retroviral AP-RT-INT alignment of consensus sequences in its "pol" ORF; Ty and copia have the AP-INT-RT alignment (Figure 1.4). It would therefore be interesting to know whether transmissible 17.6-VLPs are formed. Such a phenomenon may categorise 17.6 as an "evolutionary-link" between retroviruses and retrotransposons.

A common evolutionary ancestor for the reverse transcriptase encoding retroviruses, retrotransposons and LINEs has been proposed on the basis of consensus sequences (Fanning and Singer, 1987). Nevertheless, the very different structure of LINEs, especially at the termini of the elements (Section 1.3.1.4), could suggest that the similarities observed for these retroposons may, in fact, be the consequence of convergent, rather than divergent, evolution. The existence of a progenitor retroviral/retrotransposon element seems a more likely possibility because of the close resemblance of their genome organisations and expression strategies. Whether the "env" genes have been lost by Ty and copia or gained by retroviruses, IAP genomes and 17.6 is a matter of speculation.

10.2 Ty Elements as Selfish DNA

Pieces of DNA that confer no obvious phenotype upon the host organism, so that their only function seems to be self-perpetuation have been described as selfish DNA (Doolittle and Sapienza, 1980; Orgel and Crick, 1980). The action of natural selection on such sequences is at the DNA level itself, with respect to its ability to survive within a cell, rather than through its phenotypic effect on the host organism.

The selfish DNAs that are present in many different organisms have developed a number of characteristics that enable them to sustain themselves within their respective host cell environments. Firstly, because it does not have any effect upon phenotype, selfish DNA is able to reside within its host's genome independently of changes in the host's environment. Secondly, selfish DNA tends to be repetitive as a result of its spread by the production of additional copies of itself; this feature implies that the extremely unlikely event of multiple simultaneous deletions is required for the expulsion of the element from the host genome. The repetitive nature of selfish DNA means that it can mediate a variety of insertion and recombination events that culminate in the host's genome being somewhat fluid; thus the host may derive some selective advantage from the presence of selfish DNA within its genome. Thirdly, the propagation of selfish DNA utilises gene products that are provided by, and essential to, the host's survival (for example, proteins). If this was not so, the host could prevent the spread of selfish DNA by ceasing to supply such factors, this would not

have any deleterious effect on its own survival. Any specialised factors that are required by the selfish DNA, but not its host, are therefore provided by the selfish DNA itself.

Selfish DNA does not, however, have complete control over its own levels and mobility since the host is able to apply certain constraints upon it. Firstly, the abundance of a particular selfish DNA (repetitive element) is limited by a phenomenon known as metabolic stress. This may operate in a number of ways, for example, the fraction of the host's total energy expenditure that the selfish DNA will utilise during DNA replication, is going to be a far more significant factor for the cells of rapidly dividing unicellular organisms, than for the cells of large multicellular organisms. This fact is borne out by the far greater proportion of repetitive DNA that is present within mammalian genomes, as opposed to the genomes of single celled eukaryotes and prokaryotes (Chapter 1). It is therefore beneficial to the host (directly) and the selfish DNA (indirectly) for the amount of selfish DNA to be controlled in some way. Secondly, the expression (transcription or translation) of selfish DNA may be subject to host cell control. Consequently its integration elsewhere into the genome, the recombination events it mediates and the metabolic stress it causes may be restricted to certain stages of the host's life cycle.

The question is, therefore, whether Ty elements can be classified as selfish DNA? The answer seems to be that Ty elements are indeed selfish DNA since they possess most of the above characteristics. Firstly, Ty elements confer no obvious

phenoype to its host, S. cerevisiae; yeast strains that only have a few Ty elements still display the same phenotype as those with 30-35 (P. Philippsen, personal communication to A. Kingsman). Secondly, Ty elements are repeated sequences and contribute to both recombination (Section 1.4.2) and insertion (Section 1.4.4) events that are manifested as a variety of mutations in yeast. Thirdly, Ty elements use transcription and translation factors that are supplied by the yeast cell. The specialist functions that Ty elements require for their proposed life cycle (Figure 3.10) are encoded by the TYB gene and include an acid protease, a reverse transcriptase and possibly a specific dsDNA endonuclease. Fourthly, yeast cells do exert some control over the expression of Ty elements, as seen in the mating type regulation of Ty transcription (Section 1.4.3) and transposition (Section 1.4.6). Therefore, Ty elements may be regarded as harmless parasites of the S. cerevisiae genome.

10.3 Future Investigations

There are a number of experiments that follow the general direction of the research that has been presented and discussed in this thesis.

1. Protein requirements for VLP formation

The construction of a fine set of deletions between the endpoints of pMA5620 and pMA5620-358 followed by the subsequent particle formation assays would define exactly how much of the carboxyl terminus of p1 could be removed before particle formation ceased. Some internal deletions, within TYA, that

maintained the reading phase would also be very informative with respect to the regions of p1 that are required for the various aspects of particle formation. Furthermore, the protein domains involved in protein-protein and protein-RNA interactions could be determined.

2. RNA packaging and transposition

The demonstration that the specific signals involved in RNA packaging may be distributed throughout the genome of Ty1-15 suggests that further experiments with pMA924 and the packaging assay may be limited in their value. An alternative approach could involve the generation of a series of internal deletions within a Ty element and the examination of their effect, in the presence of a "helper" element, on both packaging and transposition (with the reservations expressed in Section 4.2). It is my expectation that increasingly larger deletions would lead to correlated decreases in the levels of packaging and in the frequencies of transposition.

3. Integrative vectors

The development of a vector system along the lines of the retroviral vectors (Figure 4.1) would be an interesting application of Ty biology. The transposition of a defective Ty element that contained a gene of interest, yet retained its ability to transpose in the presence of a "helper" element, could be used to create a stable yeast strain with chromosomally integrated copies of that gene. Growth conditions that would achieve optimal culture growth, and potentially gene expression, could then be exploited. One of the drawbacks of plasmid-borne

expression systems is the requirement for growth conditions, such as selective media, that maintain the plasmid. Alternatively, the ability of delta sequences to homologously recombine could be utilised to exchange the pre-existing chromosomal Ty elements with modified elements containing specific genes.

4. Hybrid Ty-VLPs

The insertion of ORFs into pMA562Ø and the induction of antibody responses upon immunisation with the resultant particles raises many interesting possibilities, as discussed in Section 7.5.

REFERENCES

Adams, S.E., Mellor, J., Gull, K., Sim, R.B., Tuite, M.F., Kingsman, S.M. and Kingsman, A.J. (1987) Cell **49**, 111-119.

Arrigo, S., Yun, M. and Beemon, K. (1987) Mol. Cell. Biol. **7**, 388-397.

Arkhipova, I.R., Mazo, A.M., Cherkasova, V.A., Gorelova, T.V., Schuppe, N.G. and Ilyin, Y.V. (1986) Cell **44**, 555-563.

Baltimore, D. (1970) Nature **226**, 1209-1211.

Baltimore, D. (1985) Cell **40**, 481-482.

Banerji, J., Rusconi, S. and Schaffner, W. (1981) Cell **27**, 299-308.

Beggs, J.D. (1978) Nature **275**, 104-109.

Belmont, J.W., Henkel-Tigges, J., Chang, S.M.W., Wager-Smith, K., Kellems, R.E., Dick, J.E., Magli, M.C., Phillips, R.A., Bernstein, A. and Caskey, C.T. (1986) Nature **322**, 385-387.

Bender, M.A., Palmer, T.D., Gelinas, R.E. and Miller, A.D. (1987) J. Virol. **61**, 1639-1646.

Bennink, J.R., Yewdell, J.W., Smith, G.L., Moller, C. and Moss, B. (1984) Nature **311**, 578-579.

Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) Nucl. Acids Res. **8**, 127-142.

Berg, J.M. (1986) Science **232**, 485-487.

Bishop, J.M. (1985) Cell **42**, 23-38.

Boeke, J.D., Garfinkel, D.J., Styles, C.A. and Fink, G.R. (1985) Cell **40**, 491-500.

Boeke, J.D., Styles, C.A. and Fink, G.R. (1986) Mol. Cell. Biol. **6**, 3578-3581.

Bowen, B.A., Fulton, A.M., Tuite, M.F., Kingsman, S.M. and Kingsman, A.J. (1984) Nucl. Acids Res. **12**, 1627-1640.

Bradford, M.M. (1976) Anal. Biochem. **72**, 248-254.

Broome, S. and Gilbert, W. (1985) Cell **40**, 537-546.

Brown, F., Schild, G.C. and Ada, G.L. (1986) Nature **319**, 549-550.

Brutlag, D.L. (1980) Ann. Rev. Genet. **14**, 121-144.

Byers, B. and Goetsch, L. (1975) J. Bact. **124**, 511-523.

Cameron, J.R., Loh, E.Y. and Davis, R.W. (1979) Cell **16**, 739-751.

Canaani, E., Dreazen, O., Klar, A., Rechavi, G., Ram, D., Cohen, J.B. and Givol, D. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 7118-7122.

Carlson, M. and Brutlag, D. (1978) Cell **15**, 733-742.

Cepko, C.L., Roberts, B.E. and Mulligan, R.C. (1984) Cell **37**, 1053-1062.

Chakrabiti, S., Robert-Guroff, M., Wong-Staal, F., Gallo, R.C. and Moss, B. (1986) Nature **320**, 535-537.

Chandra, P. and Steel, L.K. (1977) Biochem. J. **167**, 513-524.

Chisholm, G.E., Genbauffe, F.S. and Cooper, T.G (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 2965-2969.

Ciriacy, M. and Williamson, V.M. (1981) Mol. Gen. Genet. **182**, 159-163.

Clare, J. and Farabaugh, P. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 2829-2833.

Clark-Adams, C.D. and Winston, F. (1987) Mol. Cell. Biol. **7**, 679-686.

Clewell, D.B. and Helinski, D.R. (1969) Proc. Natl. Acad. Sci. U.S.A. **62**, 1159-1166.

Coen, E.S. and Carpenter, R. (1986) TIG **2**, 292-296.

Coen, E.S., Carpenter, R. and Martin, C. (1986) Cell **47**, 285-296.

Collins, M. and Rubin, G.M. (1983) Nature **303**, 259-260.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. and Chambon, P. (1980) Science **209**, 1406-1414.

Cryer, D.R. Eccleshall, R. and Marmur, J. (1975) in Prescott, D.M. (Ed) Methods in Cell Biology, Vol. 12, Academic Press, N.Y., pp39-44.

Cullen, B.R. (1986) Cell **46**, 973-982.

Cullen, B.R., Lomedico, P.T. and Ju, G. (1984) Nature **307**, 241-245.

Cullen, B.R., Raymond, K. and Ju, G. (1985) Mol. Cell. Biol. **5**, 438-447.

Darlix, J.-L. (1986) J. Mol. Biol. **189**, 421-434.

De Blas, A.L. and Cherwinski, H.M. (1983) Anal. Biochem. **133**, 214-219.

Delpeyroux, F., Chenciner, N., Lim, A., Malpiece, Y., Blondel, B., Crainic, R., van der Werf, S. and Streeck, R.E. (1986) Science **233**, 472-475.

DelRey, F.J., Donahue, T.F. and Fink, G.R. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 4138-4142.

Dobson, M.J., Mellor, J., Fulton, A.M., Roberts, N.A., Bowen, B.A., Kingsman, S.M. and Kingsman, A.J. (1984) EMBO J. **3**, 1115-1119.

Doolittle, W.F. and Sapienza, C. (1980) Nature **284**, 601-603.

Doring, H.-P. and Starlinger, P. (1984) Cell **39**, 253-259.

Dubois, E., Jacobs, E. and Jauniaux, J.-C. (1982) EMBO J. **1**, 1133-1139.

Duyk, G., Longiaru, M., Cobrinik, D., Kowal, R., Dehaseth, P., Skalka, A.M. and Leis, J. (1985) J. Virol. **56**, 589-599.

Eibel, H., Gafner, J., Stotz, A. and Philippsen, P. (1980) Cold Spring Harb. Symp. Quant. Biol. **45**, Part 2, 609-617.

Eibel, H. and Philippsen, P. (1984) Nature **307**, 386-388.

Elder, R.T. Loh, E.Y. and Davis, R.W. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 2432-2436.

Elder, R.T., St. John, T.P., Stinchcomb, D.T. and Davis, R.W. (1980) Cold Spring Harb. Symp. Quant. Biol. **45**, Part 2, 581-584.

Engels, W.R. (1983) Ann. Rev. Genet. **17**, 315-344.

Errede, B., Cardillo, T.S., Sherman, F., Dubois, E., Deschamps, J. and Wiame, J.-M. (1980b) Cell **22**, 427-436.

Errede, B., Cardillo, T.S., Wever, G. and Sherman, F. (1980a) Cold Spring Harb. Symp. Quant. Biol. **45**, Part 2, 593-602.

Errede, B., Cardillo, T.S., Teague, M.A. and Sherman, F. (1984) Mol. Cell. Biol. **4**, 1393-1401.

Errede, B. Company, M., Ferchak, J.D., Hutchison III, C.A. and Yarnell, W.S. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 5423-5427.

Fanning, T. and Singer, M. (1987) Nucl. Acids Res. **15**, 2251-2260.

Farabaugh, P.J. and Fink, G.R. (1980) Nature **286**, 352-356.

Fawcett, D.H., Lister, C.K., Kellett, E. and Finnegan, D.J. (1986) Cell **47**, 1007-1015.

Feinberg, M.B., Jarrett, R.F., Aldovini, A., Gallo, R.C. and Wong-Staal, F. (1986) Cell **46**, 807-817.

Fink, G.R. (1970) Methods in Enzymology **17A**, 59-78.

Fink, G.R. (1987) Cell **49**, 5-6.

Finnegan, D.J. (1985) Int. Rev. Cytol. **93**, 281-326.

Fitzgerald-Hayes, M., Clarke, L. and Carbon, J. (1982) Cell **29**, 235-244.

Flavell, A.J. and Brierley, C. (1986) Nucl. Acids Res. **14**, 3659-3669.

Flavell, A.J. and Ish-Horowicz, D. (1981) Nature **292**, 591-595.

Flavell, A.J. and Ish-Horowicz, D. (1983) Cell **34**, 415-419.

Flavell, A.J., Levis, R., Simon, M.A. and Rubin, G.M. (1981) Nucl. Acids Res. **9**, 6279-6291.

Flavell, A.J., Ruby, S.W., Toole, J.J. Roberts, B.E. and Rubin, G.M. (1980) Proc. Natl. Acad. Sci. U.S.A. **77**, 7107-7111.

Friesen, P.D., Rice, W.C., Miller, D.W. and Miller, L.K. (1986) Mol. Cell. Biol. **6**, 1599-1607.

Fulton, A.M., Mellor, J., Dobson, M.J., Chester, J., Warmington, J.R., Indge, K.J., Oliver, S.G., de la Paz, P., Wilson, W., Kingsman, A.J. and Kingsman, S.M. (1985) Nucl. Acids Res. **13**, 4097-4112.

Garfinkel, D.J., Boeke, J.D. and Fink, G.R. (1985) Cell **42**, 507-517.

Gallo, R.C. (1987) Scientific American **256**, 38-48.

Gerard, G.F., Rottman, F. and Green, M. (1974) Biochem. **13**, 1632-1641.

Gilboa, E., Mitra, S.W., Goff, S. and Baltimore, D. (1979) Cell **18**, 93-100.

Goff, S., Traktman, P. and Baltimore, D. (1981) J. Virol. **38**, 239-248.

Grandgenett, D.P., Vora, A.C., Swanstrom, R. and Olsen, J.C. (1986) J. Virol. **58**, 970-974.

Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. U.S.A. **72**, 3961-3965.

Hanahan, D. (1983) J. Mol. Biol. **166**, 557-580.

Hauber, J., Nelbock-Hochstetter, P. and Feldmann, H. (1985) Nucl. Acids Res. **13**, 2745-2758.

Haynes, J.R., Cunningham, J., von Seefried, A., Lennick, M., Garvin, R.T. and Shen, S.-H. (1986) Biotechnology **4**, 637-641.

Heide, K. and Schwick, H.G. (1978) Weir, D.M. (Ed) Handbook of Experimental Immunology, Vol. 1, 3rd Edition. Blackwell Scientific Publications, Palo Alto, California.

Hinnen, A., Hicks, J.B. and Fink, G.R. (1978) Proc. Natl. Acad. Sci. U.S.A. **75**, 1929-1933.

Hock, R.A. and Miller, A.D. (1986) Nature **320**, 275-277.

Holmes, D.S. and Quigley, M. (1981) Analyt. Biochem. **114**, 193-197.

Hwu, H.R., Roberts, J.W., Davidson, E.H. and Britten, R.J. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 3875-3879.

Jacks, T. and Varmus, H.E. (1985) Science 230, 1237-1242.

Jaenisch, R. (1983) Cell **32**, 5-6.

Jagadeeswaran, P., Forget, B.G. and Weissman, S.M. (1981) Cell **26**, 141-142.

Jat, P.S., Cepko, C.L., Mulligan, R.C. and Sharp, P.A. (1986) Mol. Cell. Biol. **6**, 1204-1217.

Jenkins N.A., Copeland, N.G., Taylor, B.A. and Lee, B.K. (1981) Nature **293**, 370-374.

Jones, K.A., Kadonaga, J.T., Luciw, P.A. and Tjian, R. (1986) Science **232**, 755-759.

Katoh, I., Yoshinaka, Y., Rein, A., Shibuya, M., Odaka, T. and Oroszlan, S. (1985) Virology **145**, 280-292.

Katz, R.A., Terry, R.W. and Skalka, A.M. (1986) J. Virol. **59**, 163-167.

Khoury, G. and Gruss, P. (1983) Cell **33**, 313-314.

Kidwell, M.G., Kidwell, J.F. and Sved, J.S. (1977) Genetics **86**, 813-833.

Kieny, M.P., Lathe, R., Drillien, R., Spehner, D., Skory, S., Schmitt, D., Wiktor, T., Koprowski, H. and Lecocq, J.P. (1984) Nature **312**, 163-166.

Kieny, M.P., Rautmann, G., Schmitt, D., Dott, K., Wain-Hobson, S., Alizon, M., Girard, M., Chamaret, S., Laurent, A., Montagnier, L. and Lecocq, J.P. (1986) Biotechnology **4**, 790-795.

Kim, S. (1986) D.Phil. Thesis. Oxford.

Kingsman, A.J., Gimlich, R.L., Clarke, L., Chinault, A.C. and Carbon, J. (1981) J. Mol. Biol. **145**, 619-632.

Kingsman, S.M. and Kingsman, A.J. (1983) Burke, D.C. and Morris, A. (Eds) Interferons, SGM Symposium, Vol. 35. Cambridge University Press, Cambridge, pp211-254.

Koyama, T., Harada, F. and Kawai, S. (1984) J. Virol. **51**, 154-162.

Kubli, E. (1986) TIG **2**, 204-209.

Kuff, E.L., Callahan, R. and Howk, R.S. (1980) J. Virol. **33**, 1211-1214.

Kuff, E.L., Feenstra, A., Lueders, K., Smith, L., Hawley, R., Hozumi, N. and Shulman, M. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 1992-1996.

Kuff, E.L., Wivel, N.A. and Lueders, K.K. (1968) Cancer Res. **28**, 2137-3248.

Laemmli, U.K. (1970) Nature **227**, 680-685.

Laski, F.A., Rio, D.C. and Rubin, G.M. (1986) Cell **44**, 7-19.

Lauer, G.D., Roberts, T.M. and Klotz, L.C. (1977) J. Mol. Biol. **114**, 507-526.

Leder,A., Swan, D., Ruddle, F., D'Eustachio, P. and Leder, P. (1981) Nature **293**, 196-200.

Lehrman, M.A., Goldstein, J.L., Russell, D.W. and Brown, M.S. (1987) Cell **48**, 827-835.

Levis, R., Collins, M. and Rubin, G.M. (1982) Cell **30**, 551-565.

Levis, R., O'Hare, K. and Rubin, G.M. (1984) Cell **38**, 471-481.

Lewin, B. (1983) Genes. John Wiley and Sons, Inc., N.Y.

Liebermann, D., Hoffman-Liebermann, B., Weinthal, J., Childs, G., Maxson, R., Mauron, A., Cohen, S.N. and Kedes, L. (1983) Nature **306**, 342-347.

Linial, M. (1987) Cell **49**, 93-102.

Little, P.F.R. (1982) Cell **28**, 683-684.

Loeb, D.D., Padgett, R.W., Hardies, S.C., Shehee, W.R., Comer, M.B., Edgell, M.H. and Hutchison III, C.A. (1986) Mol. Cell. Biol. **6**, 168-182.

Losson, R. and Lacroute, F. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 5134-5137.

Lueders, K.K., Fewell, J.W., Kuff, E.L. and Koch, T. (1984) Mol. Cell. Biol. **4**, 2128-2135.

Lueders, K.K. and Kuff, E.L. (1977) Cell **12**, 963-972.

Lueders, K.K., Segal, S. and Kuff, E.L. (1977) Cell **11**, 83-94.

Mackett, M., Smith, G.L. and Moss, B. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 7415-7419.

Maddon, P.J., Dalgleish, A.G. McDougal, J.S. Clapham, P.R., Weiss, R.A. and Axel, R. (1986) Cell **47**, 333-348.

Malim, M.H., Adams, S.E., Gull, K., Kingsman, A.J. and Kingsman, S.M. (1987b) Nucl. Acids Res. (in press).

Malim, M.H., Mellor, J.M., Wilson, W., Fulton, A.M., Kingsman, S.M. and Kingsman, A.J. (1987a) Biological Research on Industrial Yeasts, Vol. II. CRC Press, Florida (in press).

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, N.Y.

Mann, R. and Baltimore, D. (1985) J. Virol. **54**, 401-407.

Mann, R., Mulligan, R.C. and Baltimore, D. (1983) Cell **33**, 153-159.

McAleer, W.J., Buynak, E.B., Maigetter, R.Z., Wampler, D.E., Miller, W.J. and Hilleman, M.R. (1984) Nature **307**, 178-180.

McCarrey, J.R. and Thomas, K. (1987) Nature **326**, 501-505.

McClintock, B. (1950) Proc. Natl. Acad. Sci. U.S.A. **36**, 344-355.

Mellor, J., Dobson, M.J., Roberts, N.A., Kingsman, A.J. and Kingsman, S.M. (1985d) Gene **33**, 215-226.

Mellor, J., Dobson, M.J., Roberts, N.A., Tuite, M.F., Emtage, J.S., White, S., Lowe, P.A., Patel, T., Kingsman, A.J. and Kingsman, S.M. (1983) Gene **24**, 1-14.

Mellor, J., Fulton, A.M., Dobson, M.J., Roberts, N.A., Wilson, W., Kingsman, A.J. and Kingsman, S.M. (1985b) Nucl. Acids Res. **13**, 6249-6263.

Mellor, J., Fulton, S.M., Dobson, M.J., Wilson, W., Kingsman, S.M. and Kingsman, A.J. (1985a) Nature **313**, 243-246.

Mellor, J., Malim, M.H., Gull, K., Tuite, M.F., McCready, S., Dibbayawan, T., Kingsman, S.M. and Kingsman, A.J. (1985c) Nature **318**, 583-586.

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) Nucl. Acids Res. **12**, 7035-7056.

Miksicek, R., Heber, A., Schmid, W., Danesch, U., Posseckert, G., Beato, M. and Schutz, G. (1986) Cell **46**, 283-290.

Miller, A.D., Trauber, D.R. and Buttimore, C. (1986) Som. Cell and Mol. Genet. **12**, 175-183.

Miller, D.W. and Miller, L.K. (1982) Nature **299**, 562-564.

Miller, J., McLachlan, A.D. and Klug, A. (1985) EMBO J. **4**, 1609-1614.

Mossie, K.G., Young, M.W. and Varmus, H.E. (1985) J. Mol. Biol. **182**, 31-43.

Mount, S.M. and Rubin, G.M. (1985) Mol. Cell. Biol. **5**, 1630-1638.

Moss, B., Smith, G.L., Gerin, J.L. and Purcell, R.H. (1984) Nature **311**, 67-69.

Muesing, M.A., Smith, D.H. and Capon, D.J. (1987) Cell **48**, 691-701.

Murray, K., Bruce, S.A., Hinnen, A., Wingfield, P., van Erd, P.M.C.A., de Reus, A. and Schellekens, H. (1984) EMBO J. **3**, 645-650.

Nasmyth, K.A. and Reed, S.I. (1980) Proc. Natl. Acad. Sci. U.S.A. **77**, 2119-2123.

Nishizawa, M., Koyama, T. and Kawai, S. (1985) J. Virol. **55**, 881-885.

Norrander, J., Kempe, T. and Messing, J. (1983) Gene **26**, 101-106.

Norton, P.A. and Coffin, J.M. (1985) Mol. Cell. Biol. **5**, 281-290.

Ogden, J.E., Stanway, C., Kim, S., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986) Mol. Cell. Biol. **6**, 4335-4343.

O'Hare, K. and Rubin, G.M. (1983) Cell **34**, 25-35.

Ono, M., Toh, H., Miyata, T. and Awaya, T. (1985) J. Virol. **55**, 387-394.

Orgel, L.E. and Crick, F.H.C. (1980) Nature **284**, 604-607.

Pabo, C.O. and Sauer, R.T. (1984) Ann. Rev. Biochem. **53**, 293-321.

Panganiban, A.T. (1985) Cell **42**, 5-6.

Panganiban, A.T. and Temin, H.M. (1984) Cell **36**, 673-679.

Panicali, D., Davis, S.W., Weinberg, R.L. and Paoletti, E. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 5364-5368.

Paoletti, E., Lipinskas, B.R.,. Samsonoff, C., Mercer, S. and Panicali, D. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 193-197.

Paquin, C.E. and Williamson, V.M. (1984) Science **226**, 53-55.

Parkhurst, S.M. and Corces, V.G. (1985) Cell **41**, 429-437.

Parkhurst, S.M. and Corces, V.G. (1987) EMBO J. **6**, 419-424.

Paro, R., Goldberg, M.L. and Gehring, W.J. (1983) EMBO J. **2**, 853-860.

Paterson, B.M., Segal, S., Lueders, K.K. and Kuff, E.L. (1978) J. Virol. **27**, 118-126.

Patzer, E.J., Nakamura, G.R., Simonsen, C.C., Levinson, A.D. and Brands, R. (1986) J. Virol. **58**, 884-892.

Paulson, K.E. and Schmid, C.W. (1986) Nucl. Acids Res. **14**, 6145-6158.

Payne, G.S., Bishop, J.M. and Varmus, H.E. (1982) Nature **295**, 209-214.

Petes, T.D., Hereford, L.M. and Skryabin, K.G. (1978) J. Bacteriol. **134**, 295-305.

Proudfoot, N.J. and Brownlee, G.G. (1974) Nature **252**, 359-362.

Pugatsch, T. and Stacey, D.W. (1983) Virology **128**, 505-511.

Ratner, L., Haseltine, W., Patarca, R., Livak, K.J., Starcich, B., Josephs, S.F., Doran, E.R., Rafalski, J.A., Whitehorn, E.A., Baumeister, K., Ivanoff, L., Petteway Jr., S.R., Pearson, M.L., Lautenberger, J.A., Papas, T.S., Ghrayeb, J., Chang, N.T., Gallo, R.C. and Wong-Staal, F. (1985) Nature **313**, 277-284.

Rein, A., McClure, M.R., Rice, N.R., Luftig, R.B. and Schultz, A.M. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 7246-7250.

Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. **113**, 237-251.

Rio, D.C., Laski, F.A. and Rubin, G.M. (1986) Cell **44**, 21-32.

Roeder, G.S. and Fink, G.R. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 5621-5625.

Roeder, G.S. and Fink, G.R. (1983) in Shapiro, J.A. (Ed) Mobile Genetic Elements. Academic Press, N.Y., pp299-328.

Rogers, J.H. (1985) Int. Rev. Cytol. **93**, 187-279.

Rosen, C.A., Sodroski, J.G. and Haseltine, W.A. (1985) Cell **41**, 813-823.

Rubin, G.M. (1983) in Shapiro, J.A. (Ed) Mobile Genetic Elements. Academic Press, N.Y., pp329-

Saigo, K., Kugimiya, W., Matsuo, Y., Inouye, S., Yoshioka, K. and Yuki, S. (1984) Nature **312**, 659-661.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463-5467.

Schwartz, D.E., Tizard, R. and Gilbert, W. (1983) Cell **32**, 853-869.

Schwartzberg, P., Colicelli, J. and Goff, S.P. (1983) J. Virol. **46** 538-546.

Sealy, L., Moscovici, G., Moscovici, C. and Bishop, J.M. (1983) Virology **130**, 179-194.

Shampay, J., Szostak, J.W. and Blackburn, E.H. (1984) Nature **310**, 154-157.

Shank, P.R. and Linial, M. (1980) J. Virol. **36**, 450-456.

Shank, P.R., Hughes, S.H., King, H.-J., Majors, J.E., Quintrell, N., Guntaka, R.V., Bishop, J.M. and Varmus, H.E. (1978) Cell **15**, 1383-1395.

Sharp, P.A. (1983) Nature **301**, 471-472.

Shiba, T. and Saigo, K. (1983) Nature **302**, 119-124.

Shinnick, T.M., Lerner, R.A. and Sutcliffe, J.G. (1981) Nature **293**, 543-548.

Shinnick, T.M., Sutcliffe, J.G., Green, N. and Lerner, R.A. (1983) Ann. Rev. Microbiol. **37**, 425-446.

Siliciano, P.G. and Tatchell, K. (1984) Cell **37**, 969-978.

Sinclair, J.H., Burke, J.F., Ish-Horowicz, D. and Sang, J.H. (1986) EMBO J. 5, 2349-2354.

Skowronski, J. and Singer, M.F. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 6050-6054.

Smith, G.L., Godson, G.N., Nussenzweig, V., Nussenzweig, R.S., Barnwell, J. and Moss, B. (1984) Science **224**, 397-399.

Sodroski, J., Goh, W.C., Rosen, C., Dayton, A., Terwilliger, E. and Haseltine, W. (1986) Nature **321**, 412-417.

Sorge, J., Kuhl, W., West, C. and Beutler, E. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 906-909.

Sorge, J., Ricci, W. and Hughes, S.H. (1983) J. Virol. **48**, 667-675.

Spradling, A.C. and Rubin, G.M. (1982) Science **218**, 341-347.

Taguchi, A.K.W., Ciriacy, M. and Young, E.T. (1984) Mol. Cell. Biol. **4**, 61-68.

Temin, H.M. and Mizutani, S. (1970) Nature **226**, 1211-1213.

Toh, H., Hayashida, H. and Miyata, T. (1983) Nature **305**, 827-829.

Toh, H., Ono, M., Saigo, K. and Miyata, T. (1985) Nature **315**, 691-692.

Towbin, H., Staehelin, T. and Gordon, J. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 4350-4354.

Ullu, E. and Tschudi, C. (1984) Nature **312**, 171-172.

Valenzuela, P., Coit, D., Medina-Selby, M., Kuo, C.H., Van Nest, G., Burke, R.C., Bull, P., Urdea, M.S. and Graves, P.V. (1985) Biotechnology **3**, 323-326.

Valenzuela, P., Medina, A., Rutter, W.J., Ammerer, G. and Hall, B.D. (1982) Nature **298**, 347-350.

Van Beveren, C., Rands, E., Chattopadhyay, S.K., Lowry, D.R. and Verma, I.M. (1982) J. Virol. **41**, 542-556.

Varmus, H.E. (1983) in Shapiro, J.A. (Ed) Mobile Genetic Elements. Academic Press, N.Y., pp329-361.

Vogt, V.M., Wight, A. and Eisenman, R. (1979) Virology **98**, 154-167.

Wagner, M. (1986) TIG **2**, 134-137.

Walmsley, R.W., Chan, C.S.M., Tye, B.-K. and Petes, T.D. (1984) Nature **310**, 157-160.

Warmington, J.R., Waring, R.B., Newlon, C.S., Indge, K.J. and Oliver, S.G. (1985) Nucl. Acids. Res. **13**, 6679-6693.

Watanabe, S. and Temin, H.M. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 5986-5990.

Weiher, H., Konig, M. and Gruss, P. (1983) Science **219**, 626-631.

Williamson, V.M. (1983) Int. Rev. Cytol. **83**, 1-25.

Williamson, V.M., Young, E.T. and Ciriacy, M. (1981) Cell **23**, 605-614.

Wilson, I.A., Niman, H.L., Houghton, R.A., Cherenson, A.R., Connolly, M.L. and Lerner, R.A. (1984) Cell **37**, 767-778.

Wilson, S.H., Bohn, E.W., Matsukage, A., Lueders, K.K. and Kuff, E.L. (1974) Biochemistry **13**, 1087-1094.

Wilson, W., Malim, M.H., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986) Nucl. Acids Res. **14**, 7001-7016.

Winston, F., Chaleff, D.T., Valent, B. and Fink, G.R. (1984a) Genetics **107**, 179-197.

Winston, F., Durbin, K.J. and Fink, G.R. (1984b) Cell **39**, 675-682.

Winston, F. and Minehart, P.L. (1986) Nucl. Acids Res. **14**, 6885-6900.

Witte, O.N. and Baltimore, D. (1978) J. Virol. **26**, 750-761.

Yoshinaka, Y., Katoh, I., Copeland, T.D. and Oroszlan, S. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 1618-1622.

Zaret, K.S. and Sherman, F. (1982) Cell **28**, 563-573.

Zarling, J.M., Morton, W., Moran, P.A., McClure, J., Kosowski, S.G. and Hu,. S.-L. (1986) Nature **323**, 344-346.