Molecular and Genetic Studies on the *Kluyveromyces lactis* Killer Plasmids

Thesis submitted for the degree of

Master of Philosophy

at the University of Leicester

by

Simon Elliston-Elhinn

Department of Genetics

University of Leicester

July 1999

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Molecular and Genetic Studies on the *Kluyveromyces lactis* Killer Plasmids

Simon Elliston-Elhinn

Killer strains of *K. lactis* contain two linear, cytoplasmic linear dsDNA plasmids, known as k1 and k2. These plasmids have been cloned and sequenced and sequence analysis has demonstrated that these plasmids encode 14 large open reading frames. Evidence has suggested that these plasmids replicate through a protein-primed replication mechanism, similar to Adenovirus. This study has investigated the identity of some of the constituents of this replication system, specifically the relationship between the two putative DNA polymerases encoded by the k1 and k2 plasmids and the k1 and k2 terminal proteins.

Two gene hybrids, a Glutathione S-Transferase k2ORF2 and a Maltose Binding Protein-k2ORF2 were constructed and an antibody raised against the GST-k2ORF2 protein was shown to cross react with the MBP-k2ORF2 protein. Thus demonstrating the presence of anti-k2ORF2 antibodies.

ORF1 of k1 has been disrupted using a LEU2 marker and demonstrated to be an essential gene for the maintenance of the k1 plasmid. ORF1 shares significant homology to the class B (viral) DNA polymerases and this ORF is believed to encode a k1 specific DNA polymerase.

Finally, a hybrid k2ORF2-k1ORF1 DNA polymerase was constructed and used in a set of experiments which suggested that the terminal protein and DNA polymerase domains of the putative k1 and k2 DNA polymerases are not totally independent entities, as a reconstituted DNA polymerase with a k2ORF2 N-terminal domain and a k1ORF1 C-terminal domain could not replace the functionality of k2ORF2 *in vivo.*

Abbreviations

ATP	adenosine 5' triphosphate;
bp	base pairs
DNA	deoxyribonucleic acid
dNTP	deoxyribonleotide triphosphate
DTT	dithiothreitol
DPOL	DNA polymerase
DBP	DNA binding protein
ds	double stranded
EDTA	diaminoethane tetra acetic acid
IPTG	isopropyl-ß-D-thiogalactopyranoside
ITR	inverted terminal repeat
kb	kilobase pairs
kDa	kilodaltons
mRNA	messenger RNA
nt	nucleotide
ORF	open reading frame
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
TP	terminal protein
RNA	ribonucleic acid
ss	single stranded;
SSB	single stranded DNA binding protein;
SDS	sodium dodecyl sulphate;
TIR	terminal inverted repeat;
Tris	tris (hydroxymethyl) aminomethane;
UAS	upstream activating sequence;
UCS	upstream conserved sequence;
X-GAL	5-bromo-4-chloro-3-indolyl-b-D-galactoside;

Chapter 1: Introduction

1.1 Linear DNA Plasmids

Linear, extrachromosomal elements capable of autonomous replication are wide-spread in both eukaryotic and prokaryotic organisms. The first discovery of such a plasmid system resulted from investigation into cytoplasmic male sterility in maize in the 1970's. (Pring *et al*, 1977) Since then, a number of DNA plasmid systems have been discovered in organisms ranging across higher plants, yeast and gram-positive bacteria. (Esser *et al*, 1986, Meinhart *et al*, 1990, Meinhardt and Rohe, 1993). The sizes of these plasmids have, been found to range from 520-580 kb, as in the *Streptomyces parvulus* SCP1 plasmid to 1.9 kb, as for the small *Fusarium oxysporum* plasmid. The presence of characteristic Terminal Inverted Repeats (TIR's) and 5' covalently attached terminal proteins is a common structural feature which has also been identified in both bacteriophages, such as *Escherichia coli* PRD1 and ϕ29 of *Bacillus subtilis*, as well as adenovirus. (Salas, 1991) This has given rise to the hypothesis that, because of their similarity with adenovirus, linear plasmids could also utilise an adenoviral form of protein primed replication. (Meinhardt *et al* 1990, Meinhardt and Rohe, 1993, Stark *et al*, 1990)

Linear plasmids within the prokaryotic genera *Streptomyces* are particularly widespread and, in contrast to eukaryotic linear plasmids often confer a directly recognisable and often advantageous phenotype, such as antibiotic resistance, on their host organisms. (Kinashi *et al*, 1987)

Within eukaryotes the ascomycete fungi often contain linear plasmids. However, unlike in *Streptomyces* the fungal linear plasmids are often cryptic, with a few exceptions conferring a phenotype, such as the killer plasmid systems of *Kluyveromyces lactis* and the *kalilo*-plasmid induced senescence of *Neurospora crassa*.

Eukaryotic linear plasmids appear to be localised either in the cytoplasm, as is the case with yeast plasmids, or in organelles. The predominant organelle for plasmid localisation is the mitochondrion, although in the green alga *Chlamydomonas moewusii* linear plasmids have been discovered in chloroplasts. (Meinhardt and Rohe, 1993).

Table 1.1.1 Linear plasmids in yeast and Filamentous fungi (source Meinhardt *et al*, 1989)

Species	References	Plasmid Name	Size (kbp)	Localisation
Kluyveromyces lactis	Gunge *et al*, 1981 Wesolowski *et al*, 1982 Sor *et al*, 1983 Stam *et al*, 1986	pGKL1 pGKL2	8.9 13.4	Cytoplasmic Cytoplasmic
Phaffia rhodozyma	Wilber and Proffitt, 1987		8.0 2.0	
Pichia inositovora	Ligon *et al*, 1989			
Saccharomyces kluyveri	Kitada and Hishinuma, 1987	pSKL	14.2	Cytoplasmic
Saccharomyces crataegensis	Shepherd et al, 1987	pScr1 pScr2 pScr3	15.0 7.0 5.0	Cytoplasmic Cytoplasmic
Filamentous Fungi				
Agericus bitorquis	Mohan *et al*, 1984	pEM pMPJ	7.4 3.7	Mitochondrial Mitochondrial
Ascobolus immersus	Francou *et al* 1987 Meinhardt *et al*, 1986 Kempken *et al*, 1989	pA1 pAI1 pAI2 pAI3	6.4 7.9 5.6 2.8	 Mitochondrial Mitochondrial
Ceratocystis fimbriata	Giasson and Lalonde, 1987 Normand *et al*, 1987	pCF637 pFQ501	8.2 6.0	
Claviceps purpea		pClK1 pClK2 pClK3 pCCIT9 pCCIB4 pCIT5	6.7 5.5 1.1 6.7 6.7 7.1	Mitochondrial Mitochondrial Mitochondrial Mitochondrial Mitochondrial Mitochondrial
Fusarium oxysporum	Kistler and Leong, 1986	pFOXC1	8.2	
Fusarium solani	Samac and Leong, 1986	pFSC1	9.2	Mitochondrial
Morchella conica	Meinhardt and Esser, 1984	pCM31 pCM32	8.0 6.0	Mitochondrial Mitochondrial
Neurospora intermedia	Bertrand *et al*, 1985 Bertrand and Griffith, 1988	mar kalDNA	7.0 1.364	Mitochondrial Mitochondrial
Podospora anserina	Osiewacz *et al*, 1989	pAL2-1	8.4	Mitochondrial
Rhizoctonia solani	Hashiba *et al*, 1984	pRS64	2.6	

A number of the plasmids listed in the preceding table [Table 1.1.1] have been subjected to detailed analysis. The results of these analyses have shown that these linear plasmid DNAs encode RNA and/or DNA polymerases. (Kempken *et al*, 1992, Rohe *et al*, 1992) Due to this discovery it has been suggested that the linear plasmid systems rely on an intrinsic replication and transcription machinery rather than that of the host. (Meinhardt and Rohe, 1993) In order to analyse the similarity between the linear plasmids further, comparative sequence alignment of the RNA polymerases and conserved regions of the DNA polymerases of different linear plasmids, adenoviruses and bacteriophages was performed and their phylogenetic relationships deduced. (Rohe *et al*, 1992, Kempken *et al*, 1992) Data from these studies demonstrated that the DNA polymerases encoded by these extrachromosomal elements constitute a closely related group designated the type B-type viral DNA polymerases. (Bernad *et al*, 1987, Bernad *et al*, 1989)

1.2 Other Linear Plasmids

Although the linear killer plasmids pGKL1 and pGKL2 (otherwise known as k1 and k2) found in *Kluyveromyces lactis* were among the first to be discovered, protein-primed linear genetic elements have also been discovered to be widespread in yeasts, filamentous fungi and higher plants, although absent from most animal cells. (Gunge *et al*, 1981)

Throughout this range of elements emerge a number of key similarities which typify this class. The first is that all members of this family to date have been found to be non-nuclear with most linear plasmids characterised from filamentous fungi being localised to the mitochondria, whereas most of those found in yeast are cytoplasmic. A second similarity is that all linear systems encode a DNA- and RNA- polymerase. In single plasmid systems, both are encoded by a single element, whereas in multi-element systems, such as the killer plasmids of *K. lactis*, each element carries its own DNA polymerase and either encodes its own or makes use of an externally encoded RNA polymerase. Good examples of filamentous fungal and yeast linear plasmids are the *kalilo* senescence plasmid of *Neurospora intermedia* and the pSKL plasmid found in the yeast *Saccharomyces kluyveri*. The single sub-unit polymerase of k2 is unique, however, in encoding domains similar to both the *E. coli* β and β' sub-units within a single polypeptide. The relative positions of the individual conserved regions is also believed to be unique. (Wilson and Meacock, 1988)

1.2.1 The kalilo *Plasmid*

The 9.0kb *kalilo* plasmid of *Neurospora intermedia* is a linear plasmid which can induce senescence through its integration into the mitochondrial chromosome. Like other linear plasmids, the 1366 bp *kalilo* DNA molecule has a high A+T value of 70%. (Chan *et*

al, 1991) Finally, this plasmid also encodes both an RNA polymerase similar to the viral, T7 DNA-directed RNA polymerase as well a class B (viral) DNA polymerase similar to that found in the S-1 mitochondrial plasmid of *Zea mays* and other mammalian adenovirus.

1.2.2 The pSKL Plasmid

The plasmid, pSKL from the species *Saccharomyces kluyveri*, which is closely related to *Kluyvromyces lactis* is 14.2 kb in size and has a genome structure with a very high degree of similarity to one of the *K. lactis* linear plasmids. Of the 10 open reading frames analysed, all are predicted to encode proteins with above 42% amino acid sequence homology with the corresponding ORF product of the *K. lactis* linear plasmid pGKL2.

1.3 Discovery and General Biology of the *K. lactis* k1 and k2 Killer Plasmids

Shortly after the discovery of the 2 μm plasmid in *Saccharomyces cerevisiae* (Clark-Walker, 1974; Guerineau *et al*, 1971), a screen was carried out to identify other novel yeast plasmids. This involved analysing sodium dodecyl sulphate lysates of yeast protoplasts by agarose gel electrophoresis. (Gunge *et al*, 1981) Of the seventy strains analysed, belonging to seventeen different genera, one strain, *Kluyveromyces lactis* IFO1267, was found to harbour two novel plasmids. The smaller and larger plasmids were designated pGKL1 and pGKL2 respectively, or k1 and k2. These plasmids were further analysed and found to have several unusual features.

Electron-microscopic analysis revealed that plasmid fractions purified from gels consisted of homogenous linear populations of DNA molecules of 2.6-2.7 μm (5.4×10^6 molecular weight) and 4.1-4.2 μm (8.4×10^6 molecular weight) in size, thus suggesting that these plasmids were linear. (Gunge *et al*, 1981) This was confirmed by restriction enzyme analysis. Buoyant density was also measured, using neutral Caesium Chloride centrifugation, yielding a figure of 1.687 g/cm^3 for both plasmid DNAs. (Gunge *et al*, 1981) This is considerably less than both the yeast nuclear (1.699 g/cm^3) and mitochondrial (1.692 g/cm^3) DNAs suggesting that these plasmids have a very high A-T content. The plasmids were also believed to be present within the host cells at a copy number of 50-100 per plasmid/per cell. This *Kluyveromyces lactis* isolate (IFO1267) was also found to produce a toxin which could inhibit the growth of sensitive strains of *S. cerevisiae*, *Kluyveromyces lactis*, *Kluyveromyces italicus*, *Kluyveromyces thermotolerans*, and *Kluyveromyces vanudenii.* Mating experiments crossing IFO 1267 with a non-killer *Kluyveromyces lactis* strain lacking both plasmids, W600B, clearly implicated the linear plasmids in the production of this toxin, as all of the progeny from this mating both contained the linear plasmids and exhibited the "killer" phenotype and the killer plasmids. This implied that both the toxin and immunity determinants were contained on the linear plasmids as the inheritance of the killer trait showed a non-Mendelian pattern of inheritance which correlated with the pattern of inheritance of the linear plasmids. One final feature of this system was that the smaller k1 plasmid showed an absolute dependency on k2 for maintenance. The k2 plasmid, however could exist alone in the cell.

1.4 The Linear Plasmid pGKL1 (k1)

1.4.1 Genome Organisation and Nucleotide Sequence

The nucleotide sequence of plasmid k1 has been completely determined by three laboratories. Analysis revealed that the plasmid is 8874 base pairs (bp) in length. (Hishinuma *et al*, 1984; Sor and Fukuhara, 1985, Stark *et al*, 1984) [see Figure 1.4.1.1] The DNA had a uniformly high A-T content of 73.21%, which was in agreement with the buoyant density analysis. At each of the two termini of the plasmid was a copy of a 202bp Inverted Terminal Repeat (ITR). (Sor *et al*, 1985) Further analysis identified four large Open Reading Frames (ORFs), designated ORFs 1-4, or RFs 1-4, which comprised over 95% of the possible coding capacity of the plasmid. Thus, with the exception of the two 202 bp ITRs, only 50 bp of the DNA were not contained within these ORFs. Interestingly, some of the ORFs overlap, out of phase; ORF2 and ORF3 overlap on opposite coding strands and ORF4 extends into the right TIR sequence. Sequence comparisons between the predicted polypeptide products of the four open reading frames' showed no significant similarity either internally or between ORFs. Northern blot analysis of RNAs has been carried out to confirm these ORFs role as transcribed regions. (Stark *et al*, 1984) RNA extracts of *K. lactis* strains containing both linear plasmids, when probed with different fragments of k1, were found to contain three distinct transcripts consistent with the predicted sizes of the ORFs 2,3 and 4. Although we believe that ORF1 is also expressed, which I will discuss later, no ORF1 transcript could be detected, thus suggesting that this gene is only expressed at a very low level.

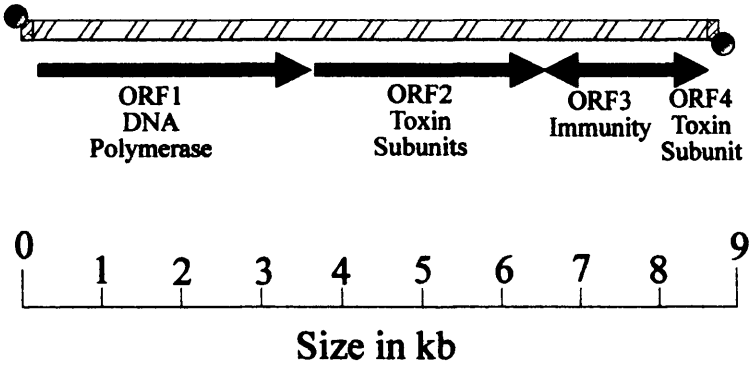


Figure 1.4.1.1: The killer plasmid pGKL1 (8874 bp) with arrows showing the extent of each of the four large open reading frames. ORF1 is the putative DNA polymerase gene, ORF2 and ORF4 are the α,β and γ sub-units of the killer toxin protein, respectively. ORF3 has been implicated as the main killer toxin immunity determinant. The terminal spheres represent the terminal proteins, covalently attached to the 5' ends of each DNA strand.

1.4.1.1 k1 Open Reading Frame 1

ORF1 is 2988 bp long and encodes a predicted polypeptide of 995 amino acids in size with a molecular mass of 116 kDa. Sequence analysis has revealed that the predicted protein product of this ORF has a high degree of similarity throughout its central and C-terminal regions to the class B family of viral DNA polymerases found in protein-primed replication systems such as PRD1, ϕ29, and Adenovirus. (Jung *et al*, 1987) Figure 1.4.1.1.1 summarises the most similar sequences as determined by a computer similarity search based on the predicted polypeptide sequence.

This search and all following similarity comparisons were done in the following way: A BLAST protein sequence similarity search was submitted for processing using the predicted amino acid sequence for the appropriate open reading frame. The predicted amino acid sequences for the best matches were then retrieved and the original and best match sequences aligned using the MACAW sequence analysis tool obtained from the EBI software database.

As would be expected from earlier analysis, k1ORF1 shows significant similarity with the Type B (protein primed) family of DNA polymerases. The level of similarity is highest throughout two of the conserved DNA polymerase domains at sequence positions 700 and 800 amino acids in to the sequence. The N-terminal sequences, in contrast do not appear to contain any conserved domains.

1.4.1.2 k1 Open Reading Frame 2

ORF2 is 3441 bp in length and encodes a predicted polypeptide of 1146 amino acids with a molecular weight of 128.95 kDa. Due to several lines of evidence, including analysis of k1 containing, non-killer deletion mutants and direct sequencing studies, this ORF has been shown to encode two of the three polypeptide components of

gure 1.4.1.1.1

chematic

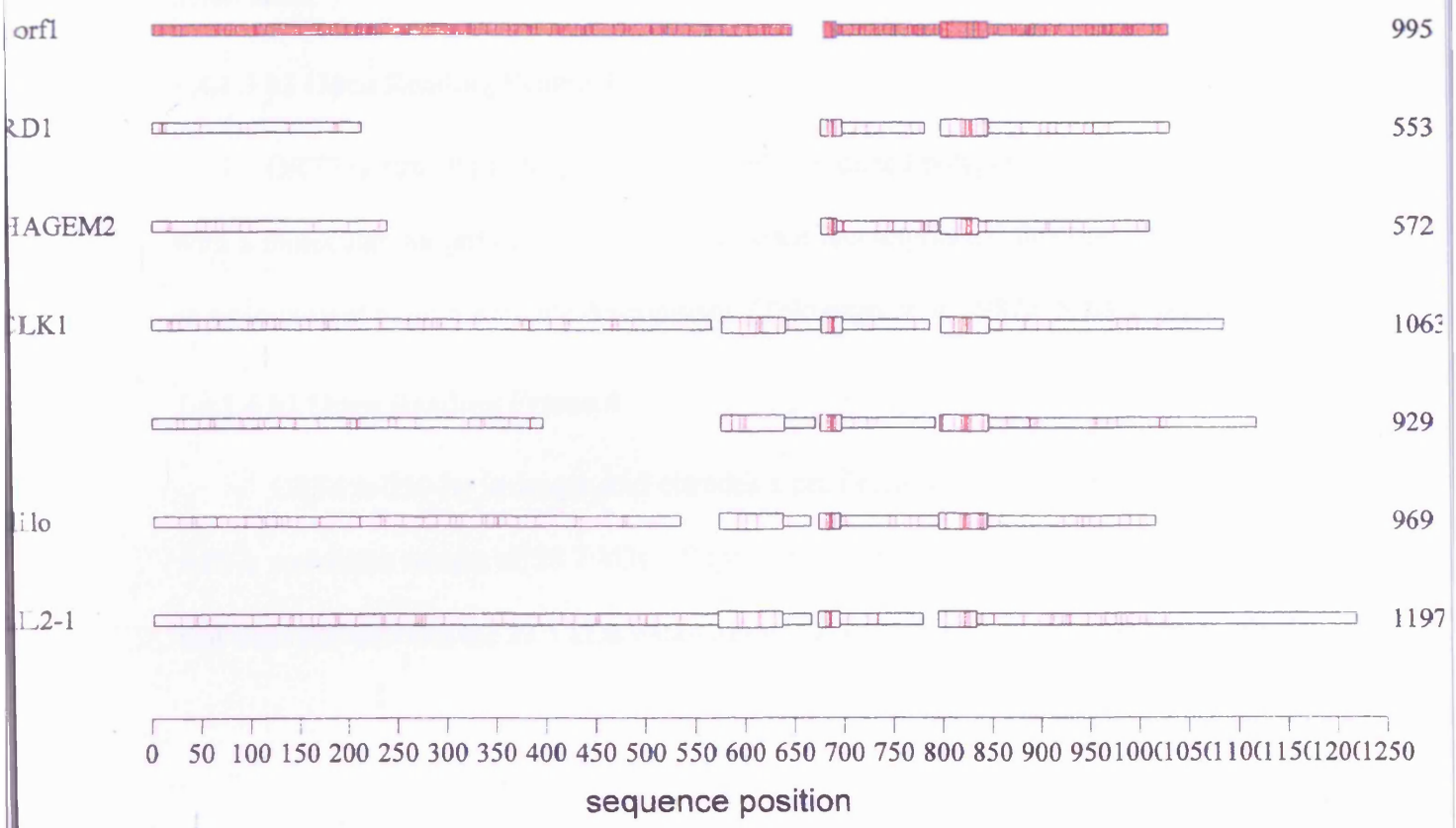


Figure 1.4.1.1.1: Similarity search results for k1ORF1. The following were the top matches found: **PRD1**-PRD1 virus DNA polymerase (protein P1) (Jung *et al*, 1987); **PHAGEM2**- Bacteriophage M2 DNA polymerase (Matsumoto *et al*, 1989); **pCLK1**- *Claviceps purpurea* mitochondrial plasmid pCKL1 possible DNA polymerase; **S1**- *Zea mays* mitochondrial plasmid S1 DNA polymerase (Oeser and Tudzynski, 1989) ; ***kalilo***- *Neurospora intermedia* mitochondrial plasmid *kalilo* probable DNA polymerase (Chan *et al*, 1991); **pAL2-1**- *Podospora anserina* mitochondrial plasmid pAL2-1 probable DNA polymerase (Hermanns and Osiewacz, 1992).
The orange highlight indicates which sequence the others were compared with. The dark lines within each sequence show small regions of similarity, with the darker segments showing higher similarity. Blocks of similar sequences are denoted by larger squares overlaying the sequence. All comparisons are done between each sequence, so sequence similarity between non-highlighted sequences is also shown.

the killer toxin, α and β. (Stark and Boyd, 1986) The killer toxin will be discussed in more detail later.

1.4.1.3 k1 Open Reading Frame 3

ORF3 is 1287 bp in length and encodes a predicted polypeptide of 428 amino acids with a molecular weight of 50.95 kDa. Evidence has implicated this open reading frame as an important toxin resistance determinant. (Tokunaga *et al*, 1987a; Stark *et al*, 1990)

1.4.1.4 k1 Open Reading Frame 4

ORF4 is 750 bp in length and encodes a predicted polypeptide of 249 amino acids with a molecular weight of 28.7 kDa. Experimental and sequencing evidence has shown that this ORF encodes the 27.5 kDa γ sub-unit of the killer toxin. (Stark and Boyd, 1986)

1.5 The Linear Plasmid pGKL2 (k2)

1.5.1 Genome Organisation and Nucleotide Sequence

The entire sequence of the k2 linear plasmid has also been elucidated. (Tommasino *et al*, 1988; Sor *et al*, 1983) Sequence analysis revealed that the DNA of plasmid k2 is 13,457 bp in length and, like k1, is extremely A-T rich (74.7%). Further analysis has discovered the presence of 10 large open reading frames which account for approximately 97% of the coding capacity of the plasmid. [see Figure 1.5.1.1] Six of the ten ORFs (ORFs: 1,4,5,6,9, and 10) are transcribed from one strand, while the other four (2,3,7 and 8) are transcribed from the other strand. There is again extensive overlapping of coding regions of the different ORFs: ORFs 8 and 9 overlap on opposite strands for the first 15, 5' codons. The start and stop codons of ORFs 1 and 2 overlap by two base-pairs. ORFs 2 and 3, 4 and 5 and 5 and 6 also overlap.

Northern analysis has also been carried out to assess which of the 10 different open reading frames are expressed. (Tommasino *et ali*, 1988) Eight differently sized species have been clearly detected, corresponding in size to ORFs 1,3,5,6,7,8,9,10. A ninth is possible, since ORF3 and ORF4 are of a similar size. This evidence suggests that open reading frames 1,3,4,5,6,7,8,9 and 10 are expressed to detectable levels. Open reading frame 2 is believed to be expressed, albeit at very low levels because it is has been shown to be an essential gene. (Schaffrath *et al*, 1995) Another feature of k2 is its inverted terminal repeats (ITRs), which are 184 bp in length.

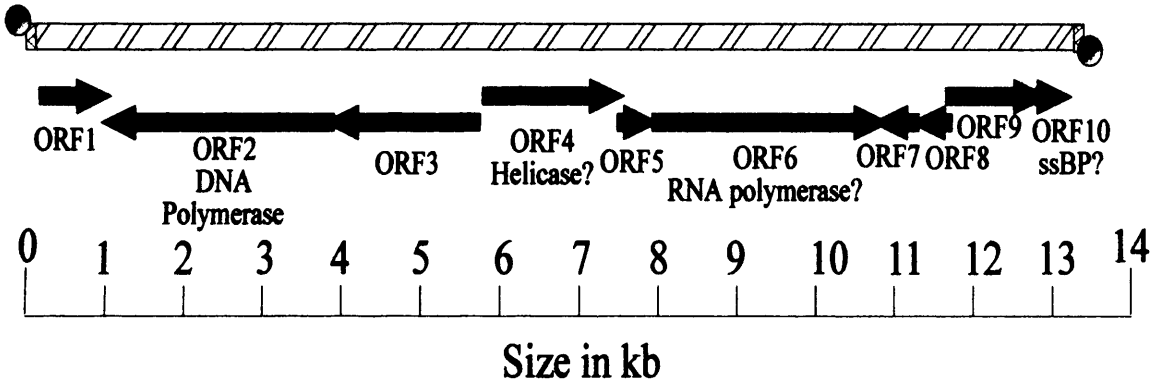


Figure 1.5.1.1: The killer plasmid pGKL2 (13457 bp)with arrows showing the extents of the large open reading frames identified. ORF1 is a non-essential gene of unknown function; ORF2 is the putative DNA polymerase gene; ORF3 is a gene of unknown function; ORF4 is an essential gene which shows similarity to the DNA helicase family of proteins; ORF5 is a gene of unknown function; ORF6 is a putative RNA polymerase gene; ORF7, 8 and 9 are genes of unknown function; ORF10 has been identified as a DNA binding gene. The spheres show the predicted locations of the terminal proteins, covalently attached to the 5' ends of each DNA strand.

1.5.1.1 k2 Open Reading Frame 1

ORF1 extends over 675 bp and encodes a protein of 224 amino acids and a molecular weight of 26.7 kDa. Although this ORF is believed to be expressed, a disruption of this open reading frame (Schaffrath *et al*, 1992) with a LEU2 marker has demonstrated that this gene product is not required for plasmid maintenance. See Figure 1.5.1.1.1 for a list of the most similar amino acid sequences. Amino acid sequence analysis was carried out as described in 1.4.1.1. Apart from the known high level of similarity between k2ORF1 and pSKL ORF1, there were no high levels of similarity detected. The best match found was that with the *E. coli* KFIC putative glycosyltransferase protein. Unfortunately KFIC has not been characterised enough to provide us with any insight though this similarity.

1.5.1.2 k2 Open Reading Frame 2

ORF2 is 2985 bp and encodes a protein of 994 amino acids with a molecular weight of 118 kDa. Computer sequence analysis has demonstrated that the predicted amino acid structure shares significant similarity with that of the Class B, viral DNA polymerases. This similarity extends through the conserved polymerase and exonuclease domains. (Tommasino *et al*, 1988) ORF2 also shares significant central and C-terminal similarity to the k1ORF1 predicted amino acid sequence. Disruption experiments have demonstrated that this gene is essential for k2 plasmid maintenance. (Schaffrath *et al*, 1995) See Figure 1.5.1.2.1 for a list of the amino acid sequences with the highest levels of similarity. As expected from earlier work, k2ORF2 showed significant similarity to the Type B (protein primed) DNA polymerases. High levels of similarity are found between the polymerase domains.

gure 1.5.1.1.1

chematic

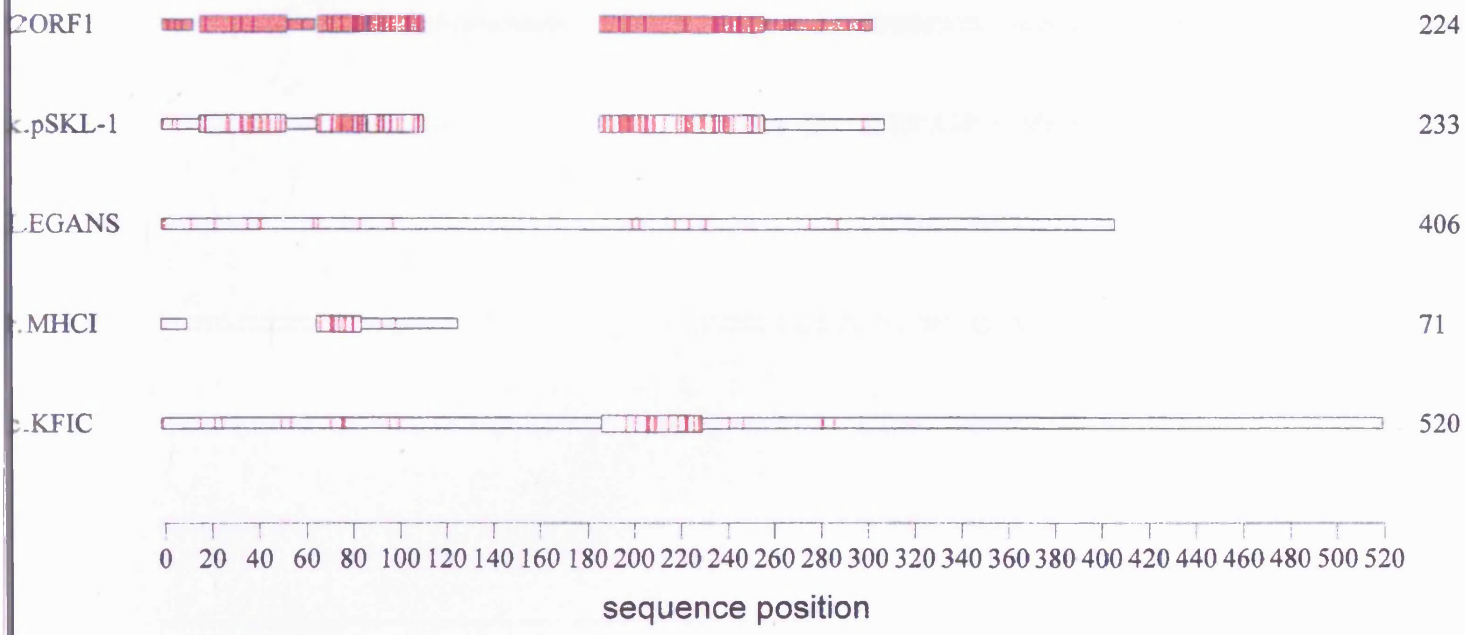


Figure 1.5.1.1.1: Similarity search results for k2ORF1. The following were the top matches found: **Sk.pSKL-1**- *Saccharomyces kluyveri* plasmid pSKL ORF1 (Hishinuma and Hirai, 1991) ; **ELEGANS**- *Caenorhabditis elegans* cosmid E02C12.6 (Wilson *et al*, 1994) ; **Pr.MHCI**- *Poecilia reticulata* MHC Class IA fragment (Sato *et al*, 1996) ; **Ec KFIC** - *Escerichia coli* KFIC gene, putative glycosyltransferase (Petit *et al*, 1995) .

igure 1.5.1.2.1

chematic

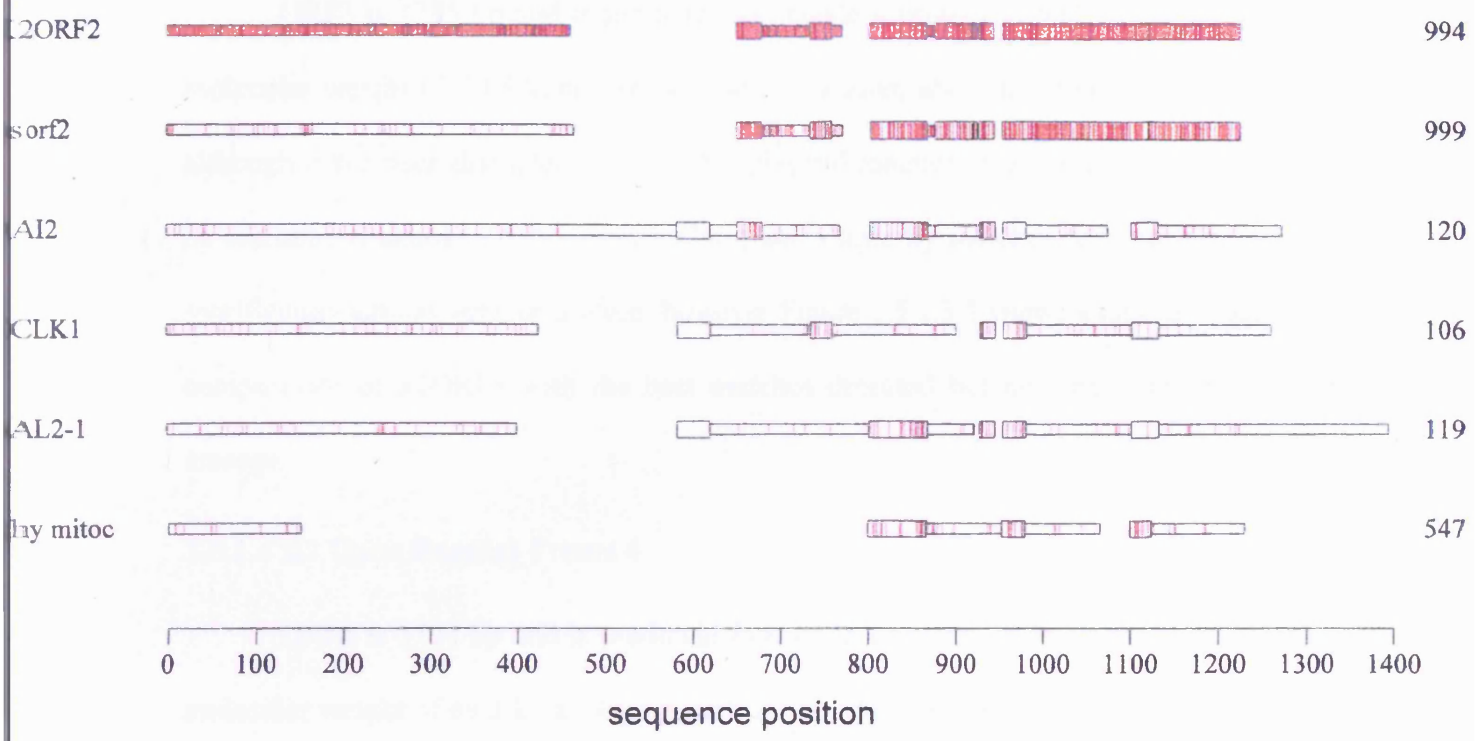


Figure 1.5.1.2.1: Similarity search results for k2ORF2. The following were the top matches found: **psorf2-** *Saccharomyces kluyveri* plasmid pSKL ORF2 (Hishinuma and Hirai, 1991); **PAI2-** *Ascobolus immersus* mitochondrial plasmid PAI2 probable DNA polymerase (Kempken *et al*, 1989); **pCLK2-1**: Podospora Anserina mitochondrial plasmid pCLK1 probable DNA polymerase (Oeser and Tudzynski, 1989); **Phy mitoc-** *Physarum polycephalum* mitochondrial DNA polymerase (Takano *et al*, 1994).

1.5.1.3 k2 Open Reading Frame 3

ORF3 is 1785 bp and is predicted to encode a protein of 594 amino acids with a molecular weight of 70.5 kDa. So far, little is known about this Open Reading Frame, although it has been disrupted to check for plasmid maintenance defect and is believed to be essential. (Fukuhara, pers. comm) Computer similarity searches have not detected any significantly similar gene or protein, however Figure 1.5.1.3.1 shows amino acid sequence comparisons of k2ORF3 with the best matches detected but no clear pattern seems to emerge.

1.5.1.4 K2 Open Reading Frame 4

ORF4 is 1794 bp and is predicted to encode a protein of 597 amino acids, with a molecular weight of 69.2 kDa. This gene has been disrupted and shown to be an essential gene. (Soond, 1994) Computer analysis shows, significant similarity between the predicted amino acid structure of this ORF product and the DNA helicases. See Figure 1.5.1.4.1 for comparison of this sequence with the most similar sequences found during a sequence similarity search.

1.5.1.5 K2 Open Reading Frame 5

ORF5 is 477 bp in length and encodes a predicted polypeptide 158 amino acids in length, with a predicted molecular weight of 18 kDa. ORF 5 has been studied in detail by R. Schaffrath. ORF5 is one of the most highly expressed k2 open reading frames in the k1/k2 system and is therefore suspected of playing an important role in the replication of k2. (Shaffrath and Meacock, 1996) Disruption analysis carried out has shown that this ORF is essential for plasmid maintenance and that it can be trans-placed onto k1 and still show functional complementation of its own deletion. This implies it is a trans-acting replication factor of unknown function. (Shaffrath and Meacock, 1996) Recently,

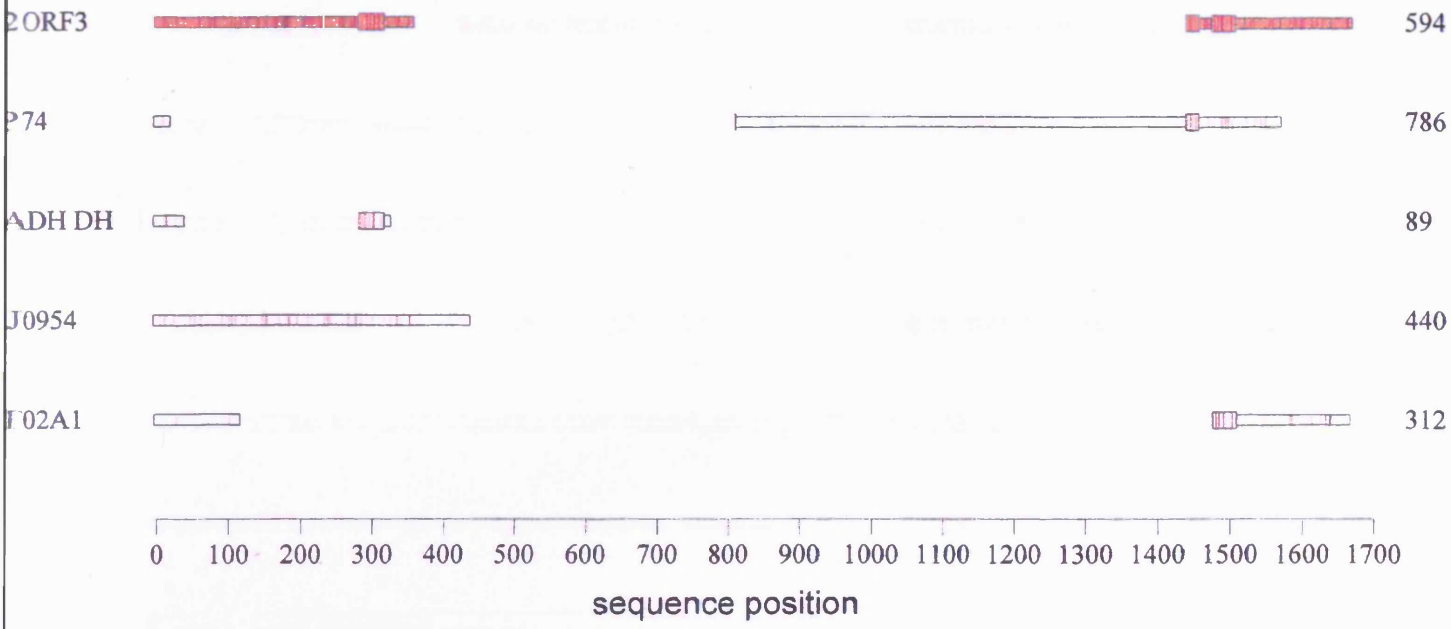


Figure 1.51.3.1: Similarity search results for k2ORF3. The following were the top matches found apart from pSKL ORF3: **BP74-** *Dictyostelium discoideum* cyclic AMP-inducible protein BP74 (Hopkinson *et al*, 1989); **NADH DH**: *Neurospora crassa* mitochondrial NADH dehydrogenase (Almasan and Mishra, 1988); **MJ0954**: *Methanococcus jannascii* hypothetical protein MJ0954 precursor (Bult *et al*, 1996) ; **RT02A1** *Acanthamoeba castellanii* RPS2 ribosomal protein S2 (Lonergan and Gray, 1993) .

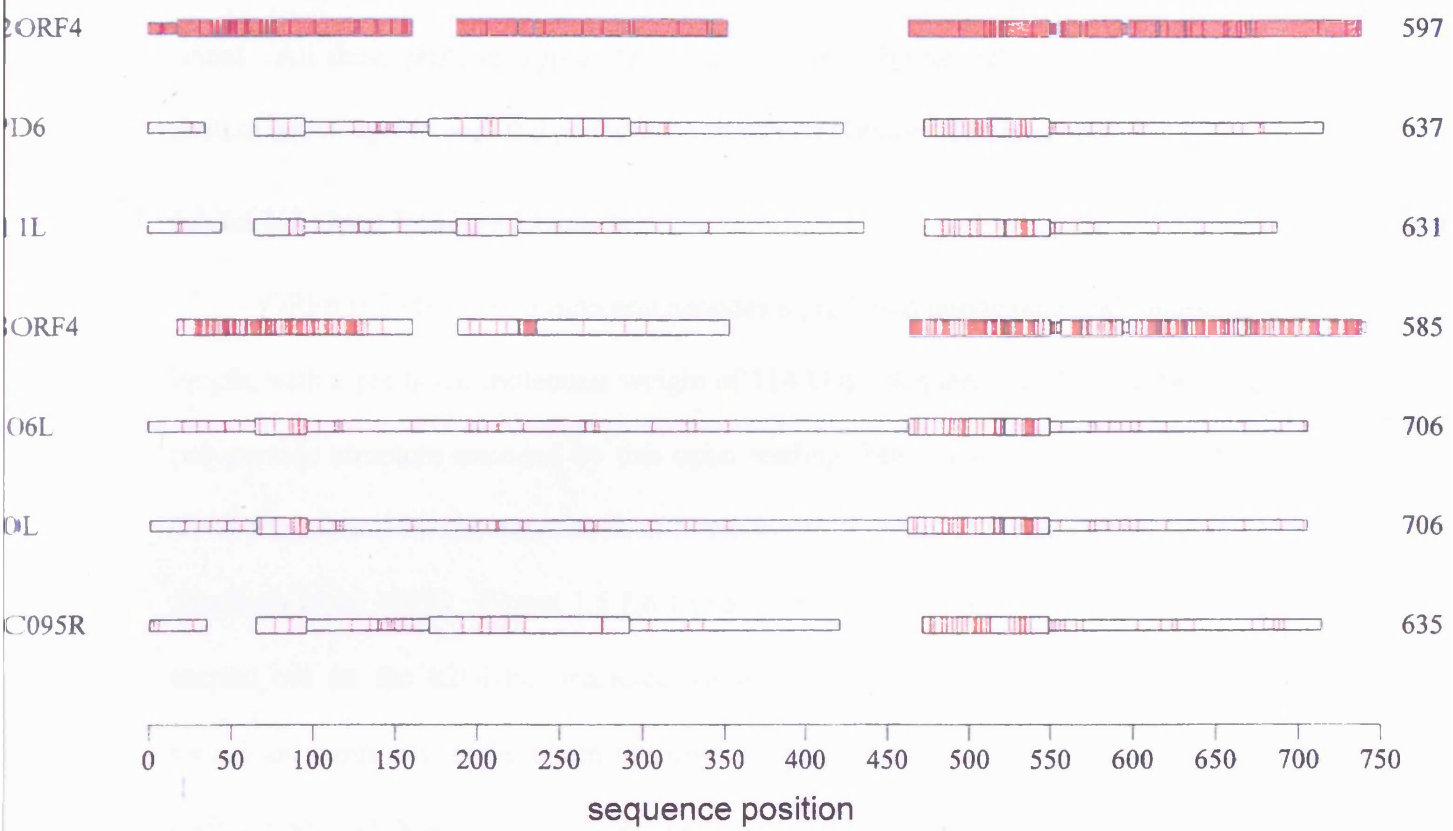


Figure 1.5.1.4.1: Similarity search results for k2ORF4. The following were the top matches found: **FPDS**- Fowlpox virus (strain FP-1) Early transcription factor 70 kDa sub-unit [potential ATP binding at amino acid position 45 and potential DEAH box motif at amino acid position 135] (Tartaglia *et al*, 1990); **D11L**- Vaccinia virus (strain Copenhagen) nucleotide triphosphatase I (Rodriguez *et al*, 1986) ; **PSORF4**: *Saccharomyces kluyveri* pSKL ORF4 (Hishinuma and Hirai, 1991) ; **8706L**- African swine fever virus helicase Q7062 (Yanez *et al*,1995); **J10L**- African swine fever helicase J10L (Vydelingum *et al*, 1993); **MC095R**- *Molluscum contagiosum* virus subtype 1 protein [contains DEAH_ATP_HELICASE motif at positions 135 - 138] (Senkevich *et al* 1996).

evidence has been gathered supporting a role as a single stranded DNA binding protein. (R. Schaffrath, pers. comm.) See Figure 1.5.1.5.1 for a list of the most similar sequences found. All these proteins appear to be some type of lysine rich, possibly DNA binding protein which agrees with the preliminary evidence obtained by Schaffrath.

1.5.1.6 K2 Open Reading Frame 6

ORF6 is 2949 bp in length and encodes a predicted polypeptide 982 amino acids in length, with a predicted molecular weight of 114 kDa. Sequence analysis of the predicted polypeptide structure encoded by this open reading frame has implicated this ORF as a RNA polymerase for the expression of linear plasmid genes. (Wilson and Meacock, 1988; Shaffrath *et al*, 1995) Figure 1.5.1.6.1 shows the results of a sequence similarity search carried out on the k2ORF6 predicted amino acid sequence. Clearly, k2ORF6 shows significant similarity to the Gram negative β and β' sub-units as well as the *Drosophila melanogaster* RNA polymerase I sub-units. Significant similarity is also shown to the 147 and 132 kDa Vaccinia RNA polymerase genes.

1.5.1.7 K2 Open Reading Frame 7

ORF7 is 399 bp in length and encodes a predicted polypeptide 132 amino acids in length, with a predicted molecular weight of 15.5 kDa. This open reading frame is believed to encode a second sub-unit of the linear plasmid RNA polymerase. (Schaffrath, 1996) Transcription and translation has been demonstrated as this ORF has been expressed in the linear plasmid system with an attached c-myc protein tag. This tag allowed the direct visualisation of this protein using Western Analysis. Disruption analysis of this ORF has suggested that this ORF is important *in cis*. (Schaffrath *et al, 1997*) Figure 1.5.1.7.1 shows comparisons of the K2ORF7 amino acid sequence against the

g ure 1.5.1.5.1

chematic

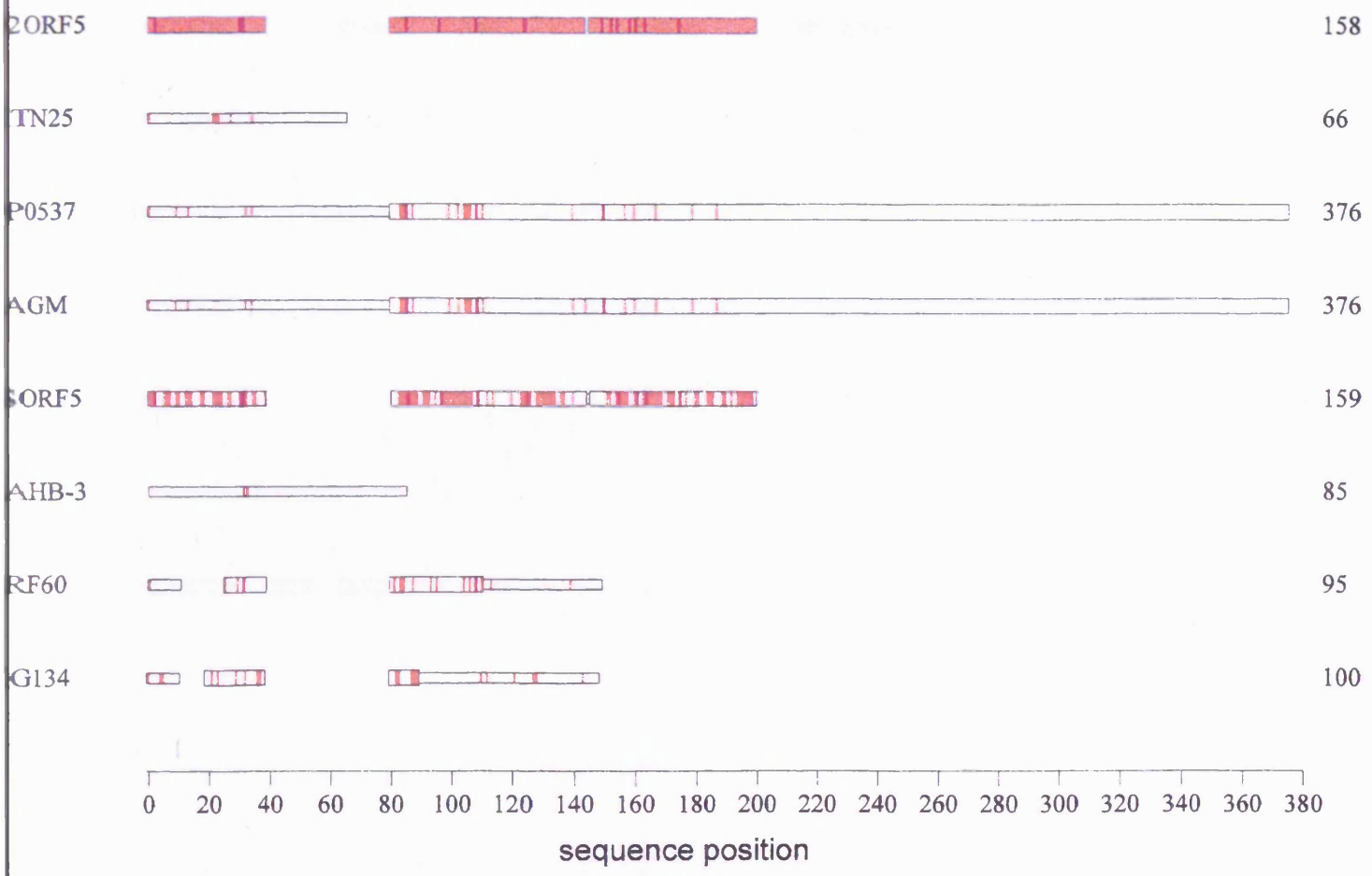


Figure 1.5.1.5.1: Similarity search results for k2ORF5. The following were the top matches found: **MTN25-** *Meicago truncatula* protein MTN25 [lysine rich] (Gamas *et al*,1996); **HP0537-** *Helicobacter pylori* CAG pathogenicity island protein CAG16 (Tomb *et al*, 1997); **CAGM-** *Helicobacter pylori* CAGM gene (Rappuoli and Covacci, 1996); **PSORF5-** *Saccharomyces kluyveri* pSKL ORF5 (Hinshinuma and Hirai, 1991); **HAHB-3-** Homeodomain containing protein from *Helianthus Annuus* (Chan and Gonzalez, 1994)**; ORF60-** *Lymantria dispar* multicapsid nuclear polyhedrosis virus hypothetical 10.9 kDa protein in LEF8-FP intergenic region (Bischoff and Slavicek, 1996); **MG134-** *Mycoplasma genitalium* hypothetical protein MG134 (Fraser *et al*, 1995).

igure 1.5.1.6.1

Schematic

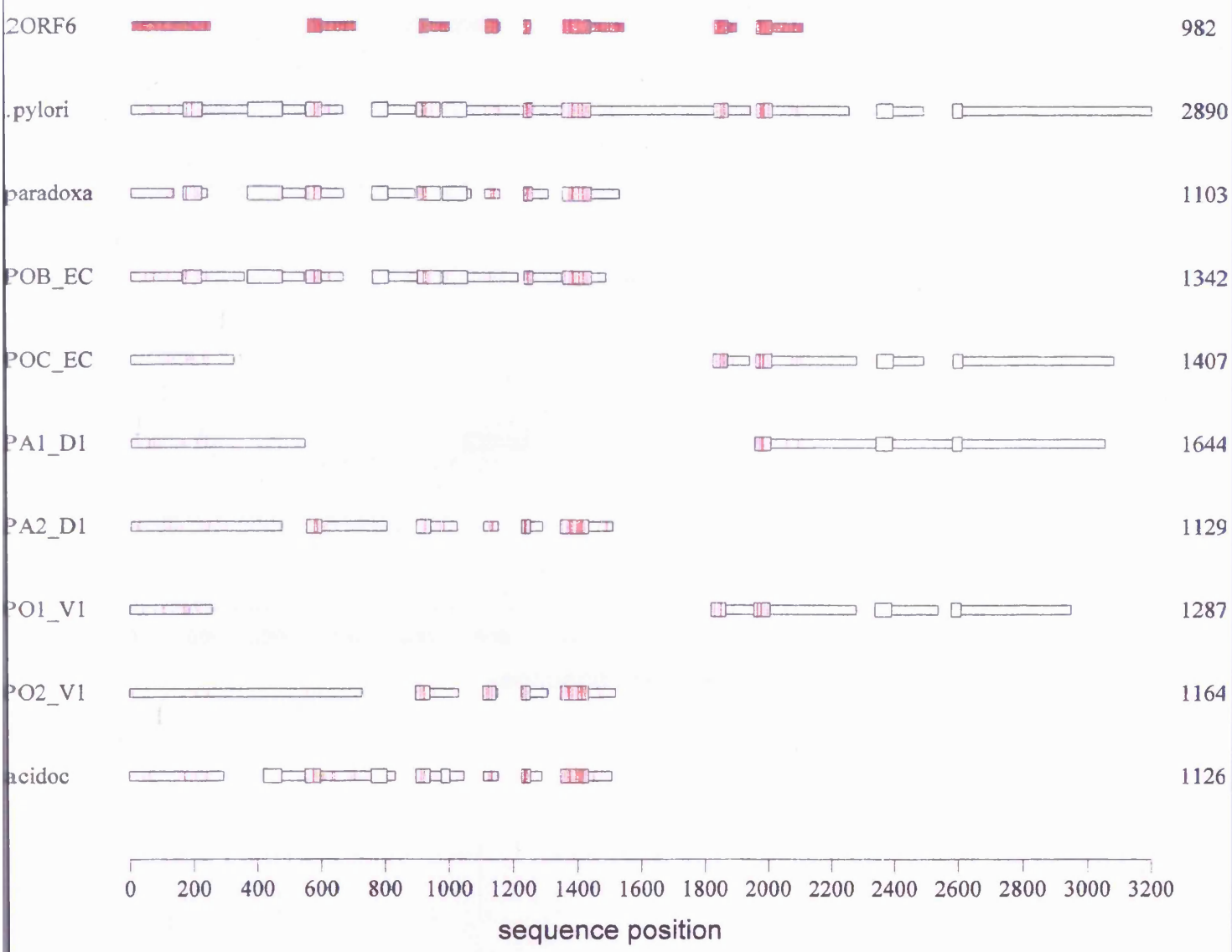


Figure 1.5.1.6.1: Similarity search results for k2ORF6. ***H. pylori***- *Helicobacter pylori* DNA directed RNA polymerase beta sub-unit (Tomb *et al*, 1997); ***C paradoxa***- *Cyanphora paradoxa* DNA directed RNA polymerase beta sub-unit (Stirewalt, unpublished)**; RPOB_EB**- *Escherichia coli* DNA directed RNA polymerase beta sub-unit (Ovchinnikov *et al*, 1981); **RPOC_EC**- *E. coli* DNA-directed RNA polymerase β' chain(Ovchinnikov *et al*, 1982); **RPA1_DM**- *Drosophila melanogaster* DNA-directed RNA Polymerase I largest subunit (Knackmuss *et al* unpublished); **RPO1_VV**- Vaccinia virus DNA-directed RNA polymerase 147 KD polypeptide (Broyles and Moss, 1986); **RPO2_VV**- Vaccinia virus DNA-directed RNA polymerase 132 KD polypeptide (Amegadzie *et al*, 1991);***S acidoc***- *Sulfolobus acidocaldarius* DNA directed RNA polymerase beta sub-unit (Puehler *et al*, 1989).

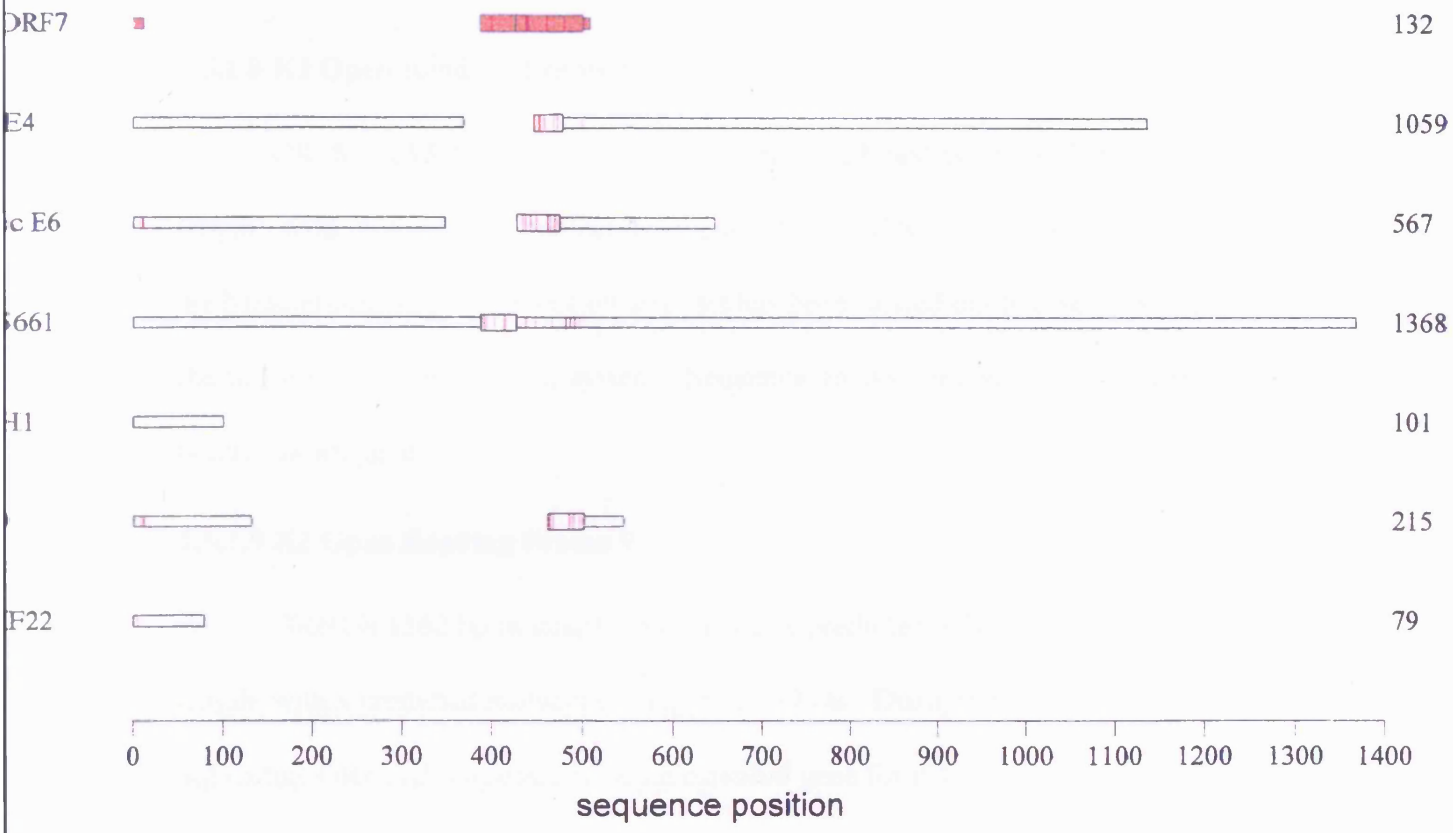


Figure 1.5.1.7.1: Similarity search results for k2ORF7. The following were the top matches found excluding pSKL plasmid: **DRE4-** *Drosophila melanogaster* probable transcriptional regulator (Washington *et al*, unpublished); **Vacc E6-** Vaccinia virus protein E6 (Gershon *et al*, unpublished); **Q25661-** *Plasmodium chabaudi* putative glutamate and asparagine rich protein (Wiser and Giraldo, unpublished); **QPH1-** QPH1 plasmid ORF100 fragment *Coxiella burnetti* (Thiele *et al*, 1995); **C19-** Swinepox virus protein C19 (Massung *et al*, 1993); **YMF22-** *Marchantia polymorpha* hypothetical 9.3 kDa protein in NAD3-NAD7 intergenic region [mitochondrial] (Oda *et al*, 1992)

closest matches obtained from a BLAST sequence similarity search. No very significant matches were found to any known protein.

1.5.1.8 K2 Open Reading Frame 8

ORF8 is 288 bp in length and encodes a predicted polypeptide 95 amino acids in length, with a predicted molecular weigh of 11 kDa. This ORF has so far been uncharacterised, and no disruption analysis has been carried out to assess its importance to the linear plasmid replication system. Sequence analysis reveals no significant similarity to any known gene.

1.5.1.9 K2 Open Reading Frame 9

ORF9 is 1362 bp in length and encodes a predicted polypeptide 453 amino acids in length, with a predicted molecular weight of 55 kDa. Disruption analysis has been carried out on this ORF and it appears to be an essential gene for plasmid maintenance. (Fukuhara, pers. comm.) This ORF has no significant sequence similarity to any other known gene. See Figure 1.5.1.9.1 for best similarity search matches. Of those proteins detected in this search no clear pattern emerged to provide any significant hint as to what the function of this protein is.

1.5.1.10 K2 Open Reading Frame 10

ORF is 312 bp in length and encodes a predicted polypeptide 103 amino acids in length, with a predicted molecular weight of 12 kDa. Tommasino reported that the ORF10 protein product successfully bound DNA-Sepharose and could be eluted with 350 mM salt. (Tommasino, 1991) This ORF has also been overexpressed in *Escherichia coli* and the overexpressed protein has been found to specifically bind the Inverted Terminal Repeat (ITR) regions of k2. (McNeel and Tamanoi, 1991) This gene was hence designated TRF1 (Terminal Recognition Factor 1). While TRF1 is known to bind the termini k2, it is not

Figure 1.5.1.9.1

Schematic

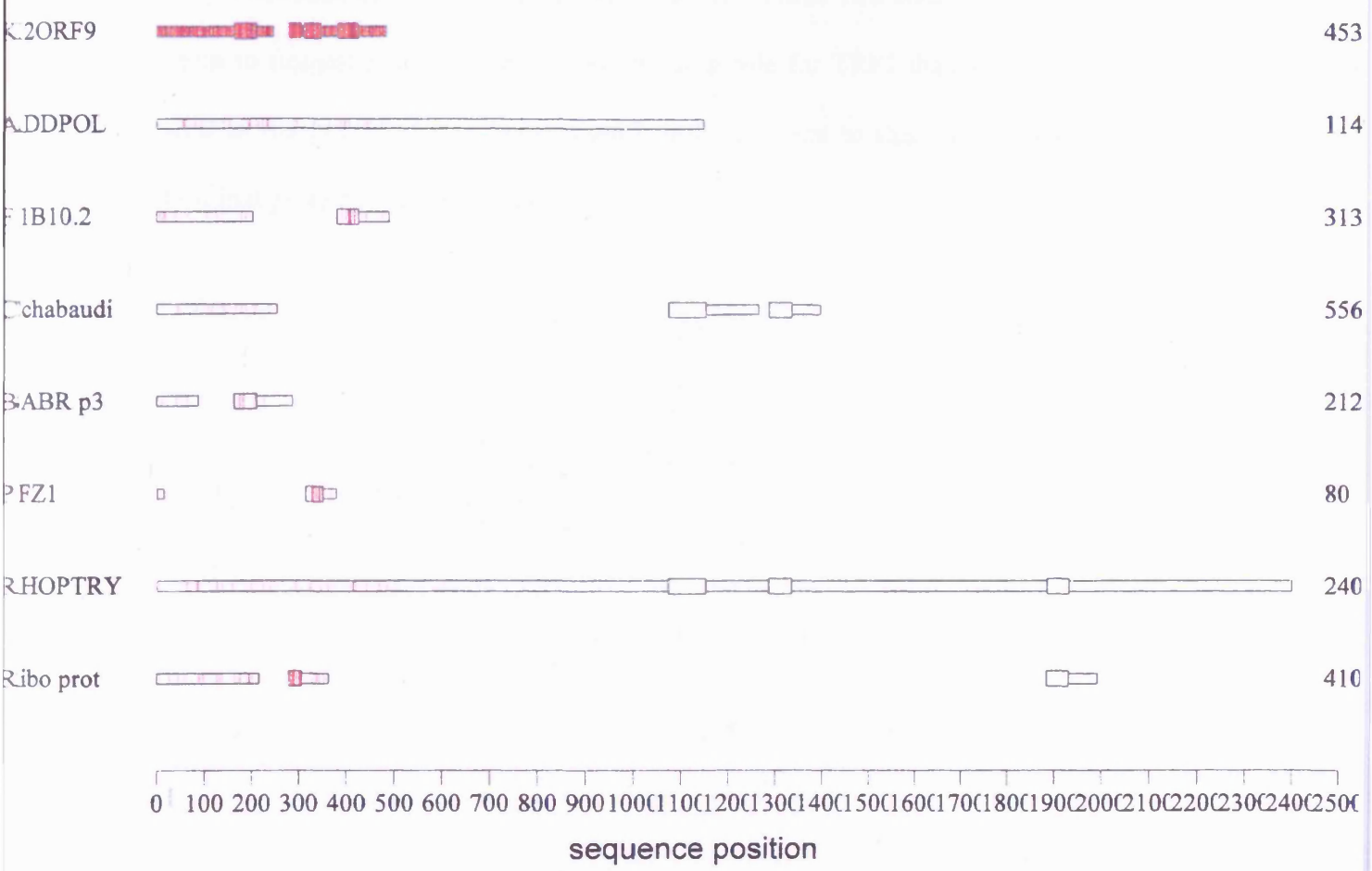


Figure 1.5.1.9.1: Similarity search results for k2ORF9. The following were the top matches found excluding pSKL plasmid: **ADDPOL-** Canine adenovirus 1 DNA polymerase (Campbell and Zhao, unpublished); **F1B10.2-** *Caenorhabaitis elegans* protein similar to heavy chain myosin (Wilson *et al*, 1994); **Cchabaudi-** *Chabaudi chabaudi* NE-rich protein (Snounou *et al*, 1988); **BABR p3-** *Babesia bovis* gene BABR protein 3 (Cowman *et al*, 1984); **PFZ1-** *Methanobacterium thermoformicium* plasmid pFZ1 (Noelling *et al*, 1992); **RHOPTRY-** *Bergei yoelii* rhoptry protein (Sinhal *et al*, 1996); **RIBO PROT-** *Emericella nidulans* mitochondrial protein S5 (essential for mitochondrial protein synthesis and required for the maturation of small ribosomal sub-units) (Lazarus and Kuntzel, unpublished)

believed that it is the terminal protein, from which plasmid replication initiates, because the interaction is not believed to be covalent. These two results taken together would seem to suggest a more general DNA binding role for TRF1 than simply ITR binding as well as the predicted size being significantly different to that shown by preparations of terminal proteins. (Reay,1997)

1.6 Protein Primed Replication

One of the most fascinating aspects of the k1-k2 linear killer plasmid system is its DNA replication which remains, as yet, largely unelucidated. Sequence analysis has suggested that both k1 and k2 encode their own DNA polymerases which could, presumably, be used to replicate the plasmids in the absence of the nuclear DNA polymerases of the host organism. (Tommasino *et al*, 1988; Jung *et al*, 1987) The use of their own DNA polymerases to drive replication, as well as the known linear nature of these plasmids and the presence of covalently attached terminal proteins suggested that the linear killer plasmids use a "protein- primed" replication system.

1.6.1 Evidence for Terminal Proteins

Further study of the general biology and structure of the killer plasmid DNA molecules themselves has demonstrated that the termini of the linear plasmids appear to be "blocked"; by a protein covalently attached to the 5' end of the DNA strand. Several lines of evidence suggest this. The first was that attempts to radio-label the 5' ends of the plasmid DNA with [γ] ^{32}P-ATP were unsuccessful. Killer plasmid DNAs were also resistant to the 5'-3' exonuclease activity of the λ - exonuclease but were sensitive to the 3'-5' exonuclease activity of Exonuclease III. (Stam *et al*, 1986) Experimental evidence has further defined the sizes of the k1 and k2 terminal proteins as 28 kDa and 36 kDa respectively. This was demonstrated through two methods. The first was a set of experiments of carried out by Stam *et al* who demonstrated that terminal restriction fragments of k1 and k2 with attached terminal proteins did not enter an agarose gel matrix. Also ^{125}I labelling of restriction fragments followed by agarose electrophoretic analysis of these restriction fragments also provided good evidence for the presence of terminal

proteins. (Stam *et al*, 1986) Further work by Reay in which the k1 and k2 terminal proteins were separated from killer plasmids, labelled with ^{125}I and subjected to polyacrylamide electrophoresis confirmed the sizes predicted by Stam *et al.* (Reay,1997)

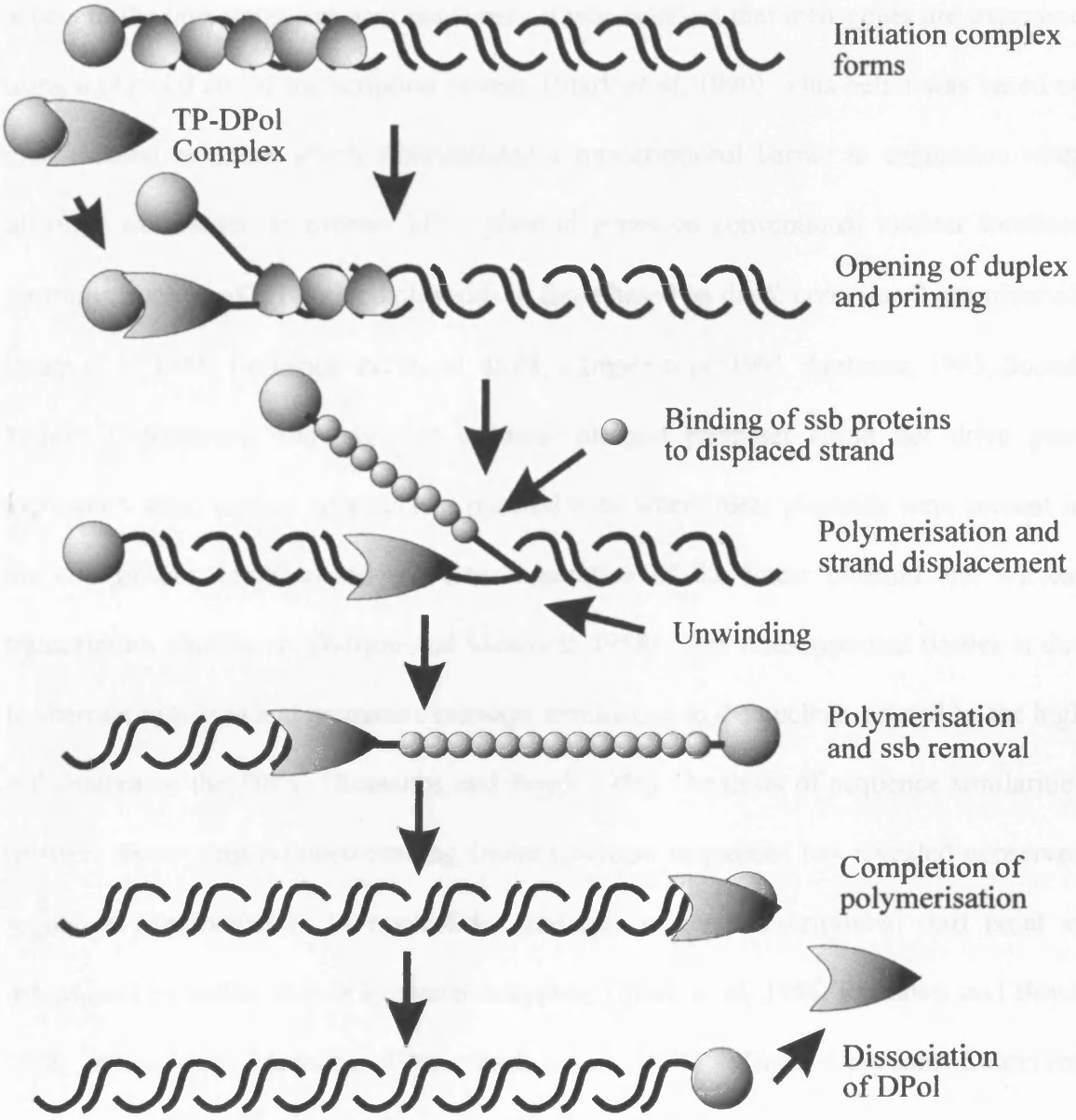
1.6.2 Functions of Terminal Proteins

The fact that these plasmids are linear and have covalently attached terminal proteins suggests several things about their mode of replication. The first is that, since they do not appear to have telomeres, they must have some way to overcome the continual 5' deletion of DNA that would result from the use of RNA primers for replication. One method used to overcome this problem is for linear molecules to initiate replication from a protein bound to the 5' end of the DNA molecule, this is known as protein priming and all known examples of this utilise a covalently attached terminal protein. Well characterised examples of this include the animal virus, Adenovirus, and the bacteriophages PRD1 which infects *E. coli* and *S. typhimurium*, and ϕ29 which infects *Bacillus subtilis.* (Bamford *et al*, 1983; Challberg *et al*, 1982; Watabe and Ito, 1983) Experimental evidence from Adenovirus and bacteriophages has demonstrated the importance of the covalently attached terminal protein in terminus recognition by the initiation complex. (Caldentey *et al*, 1992; Challberg and Rawlins, 1984) This suggests that the terminal protein can recognise itself either alone or when in a complex with other components. In addition to this role in DNA replication, terminal proteins presumably perform several other functions. One of the key functions could be to protect the termini of the DNA molecules from 5'-exonuclease attack and recombination within the cell.

1.6.3 Protein Primed Replication

Protein primed replication initiation occurs in a similar form in several different organisms. Essentially what happens is that an interaction between the terminal protein, DNA polymerase, a specific dNTP and accessory factors, such as DNA binding proteins (DBP) and single-stranded binding proteins, forms an initiation complex. This contains a terminal protein molecule attached to a specific dNTP through an ester linkage to a serine, threonine or tyrosine residue of the terminal protein. This charged terminal protein and other factors, bind and recognise the terminal region of the DNA molecule to be replicated as well as its associated terminal protein. An initiation complex is formed at the extreme terminus of the molecule, the 5'-3' strand displaced and the specific dNTP attached to the terminal protein forms a stable base pair with the extreme 3' base on the complementary template. The DNA polymerase then catalyses new strand formation and the parent strand is displaced, with single-strand DNA binding proteins stabilising the displaced complementary strand. There are two theories as to how the second strand is synthesised. The first is that the replication is somehow co-ordinated over the two strands and leads to the simultaneous formation of the two daughter plasmids. The second is that one daughter plasmid is produced, displacing the other strand during replication and that the displaced strand forms a "pan-handle" structure which is then replicated by a separate replication initiation event. [see Figure 1.6.3.1] In the example of ϕ29 the DNA polymerase catalyses the formation of the terminal protein-deoxynucleotide link. Interestingly, in the absence of ϕ29 DNA to guide the deoxynucleotidation a formation can be formed between any dNTP and the terminal protein. This is in contrast to the situation when the DNA is present, when a strong dATP preference is shown. (Blanco *et al*, 1992)

Figure 1.6.3.1: Model for Protein-Primed Replication of DNA
(Based on Salas, 1991)



1.7 Expression of Killer Plasmid Open Reading Frames

1.7.1 Transcription

Due to the cytoplasmic nature of the *K. lactis* killer plasmids which denies them access to the host transcriptional machinery it was believed that their genes are expressed using a plasmid coded transcription system. (Stark *et al*, 1990) This belief was based on experimental evidence which demonstrated a transcriptional barrier to expression when attempts were made to express killer plasmid genes on conventional nuclear localised plasmids, such as pKD1 derived plasmids or those based on the *S. cerevisae* 2 μm plasmid. (Stam *et al*, 1986, Romanos and Boyd, 1988, Kämper *et al*, 1991, Ambrose, 1993, Soond, 1994) Experiments showing that a linear plasmid promoter could not drive gene expression when carried on a nuclear plasmid even when linear plasmids were present in the cell further demonstrated the clear separation of the linear plasmid and nuclear transcription machinery. (Wilson and Meacock, 1988) This transcriptional barrier is due to aberrant initiation and premature message termination in the nucleus, caused by the high AT content of the DNA. (Romanos and Boyd, 1988) Analysis of sequence similarities between linear plasmid open reading frame upstream sequences has revealed conserved sequences approximately 14 nucleotides upstream of the transcriptional start point as determined by earlier primer extension mapping. (Stark *et al*, 1984, Romanos and Boyd, 1988, Wilson and Meacock, 1988, Stark *et al*, 1990) These Upstream Conserved Sequences (UCS) have a general form of ACT $^{A}/_{T}$ AATATGA.

The importance of these UCS's in gene expression has been demonstrated by their ability to drive expression of nuclear genes when fused at an appropriate distance upstream

of the transcriptional start site of the gene, in contrast to the failure of conventional yeast promoter elements to drive transcription in killer plasmids. (Kämper *et al*, 1989, Kämper *et al*, 1991, Gallo and Galeotti, 1990, Tanguy-Rogeau *et al*, 1990, Schaffrath *et al*, 1992) Successful transcription of nuclear genes inserted into k1 arises through plasmid circularisation or addition of telomeres to the k1 plasmid and its re-localisation into the nucleus and replication through the use of fortuitous ARS sequences. (deLouverncourt *et al*, 1983, Kämper *et al*, 1989, Kämper *et al*, 1991, Meinhardt *et al*, 1990) All these data clearly indicated that a separate plasmid encoded transcription system is present in the cell. (Romanos and Boyd, 1988, Wilson and Meacock, 1988, Stark *et al*, 1990)

The next question to be resolved was, what were the constituent parts of the transcription system. The k2ORF6 product has been implicated as the probable RNA polymerase for both the k1 and k2 plasmids. (Wilson and Meacock, 1988, Stark *et al*, 1990) Sequence analysis has shown that the 974 amino acid polypeptide shares amino acid sequence homology to both the β and β' sub-units of the *E. coli* RNA polymerase, encoded by the *rpoB* and *rpoC* genes, respectively as well as the corresponding sub-units of the RNA polymerase of the eukaryotic RNA Pol II homologues, also *Nicotiana tabacum* chloroplasts and the liverwort *Marchantia polymorpha.* (Wilson and Meacock, 1988) Sequence analysis on ORF7 has also suggested it as a possible component in the transcription apparatus due to its sequence homology to a conserved carboxy-terminal domain of the *E.coli rpoC(β')* which is missing from ORF6. (Rowland, Glass and Meacock, unpublished data; Schaffrath, PhD thesis)

A third possible component could be the k2 ORF4 gene product. Which shows significant homology to a vaccinia virus DNA-dependent ATPase. (Wilson and Meacock, 1988)

The k1 transcripts have 20 - 50 nucleotide 5' untranslated leader regions, whereas k2 has untranscribed 5' leader regions of between 30 and 110 nucleotides. The transcripts are also believed to not be poly-adenylated as they are excluded, for the most part, from the nuclear transcription machinery. Sequence analysis data published by Meinhart *et al* has also demonstrated that K2ORF3 may encode a viral-like RNA capping enzyme. (Larsen *et al*, 1998)

1.7.2 Translation

Several features might affect the rate of the translation of linear plasmid transcription products. Firstly is that it is possible that the mRNA produced by linear plasmids is uncapped. This would have a detrimental effect on the translation of these mRNA in relation to the other, capped mRNAs competing for the same 80s ribosomes *in vivo*, because of the importance of this capped region in stimulating the efficiency of ribosome binding. (Kozak, 1983) This lack of capping could require the presence of additional factors to stimulate translation, possibly, produced by the linear plasmids.

One example of adaptation of uncapped RNA is *S. cerevisiae* VLP's which, it has been suggested, have sequences within their 5' region which allow them to hybridise with the 3' end of the 5.8S or 18S ribosomal RNA thus stimulating binding efficiency. (Leibowitz *et al*, 1983) The diminutive 5' untranslated regions could also have an effect. The third factor which could have a major effect on translational efficiency is the very high A+T content which, through the constraints it places on codon usage and hence bias in the tRNA population used, slow down translation.

1.8 The Killer Phenotype

1.8.1 Toxin Structure

Strains of *Kluyveromyces lactis* harbouring the k1 and k2 linear plasmids exhibit a "killer" phenotype, hence the name "killer plasmids". This phenotype exhibits itself as an inhibition of sensitive strain growth surrounding a "killer" *K. lactis* colony. This killer phenotype is mediated through the action of a large secreted hetero-trimeric glycoprotein toxin. The killer toxin was first purified by Stark and Boyd and demonstrated to consist of three sub-units, designated α (97 kDa), β (31 kDa), and γ (28 kDa). (Stark and Boyd, 1986) Further analysis demonstrated that secretion of this toxin is inhibited by tunicamycin, a glycosylation inhibitor, thus suggesting that the killer toxin is a glycoprotein. The open reading frames encoding the toxin sub-units have been identified, through sub-unit purification followed by amino acid sequence analysis, to be k1 ORF2 for the α and β sub-units and k1 ORF4 for the γ sub-unit. (reviewed by Stark *et al*, 1990)

1.8.2 Toxin Biosynthesis

Following the transcription and translation of the k1 ORF2 and k1 ORF4 genes polypeptide processing occurs. The primary translation products have the typical amino-terminal hydrophobic regions of secreted proteins and hence are passed into the endoplasmic reticulum where the leader sequences are removed by signal peptidase. Proteolytic cleavage is required for the separation of the initial ORF2 peptide into the α and β sub-units. This proteolytic separation of the α and β sub-units has been shown to require a *KEX* like protein, contained within the Golgi, and a single mutation, *kex1* has

been found to prevent toxin secretion in *K. lactis*. Upon the cloning by complementation and sequencing of the wild-type *KEX1* gene, it was found to be a serine protease which shares structural similarities to the *S. cerevisiae KEX2* gene. (Wésolowski-Louvel *et al*, 1982) (review in Stark *et al*, 1990)

1.8.3 Toxin Secretion

1.8.3.1 Mode of Action

The toxin secreted by $k1^+$ $k2^+$ strains of *Kluyveromyces lactis* can inhibit the growth of a surrounding lawn of sensitive *K. lactis*. It has been shown to have similar inhibitory effects on several other yeast strains, including *Kluyveromyces thermotolerans, S. cerevisiae, Toruplopsis glabrata, Candida utilis, Candida intermedia, Kluyveromyces vanudenii, S. italicus, S rouxii*, but has no effect on some species of yeast, including *Schizosaccharomyces pombe*. The effects and patterns of toxin secretion have been studied in some depth, although there is no definitive answer as to the utility of this toxin in the wild or as to its mode of action. Such work has also shown that diploid strains show a marked decrease (20 - 100 fold) in sensitivity in comparison with their haploid derivatives. Data obtained for the effectiveness of toxin inhibition of cell growth following release from Hydroxyurea induced S-phase arrest suggests that the toxin is principally active at G1. Several toxin resistant mutants have been isolated, some of which have helped clarify the situation. (Gunge *et al*, 1981; Tokunaga *et al*, 1989; Butler *et al*, 1991; Butler *et al*, 1994)

Toxin production has been demonstrated to be tightly coupled to vegetative growth and reach its maximum production in late exponential phase. Stark *et al* have demonstrated

that the toxin shows maximal activity at pH 7.0 and has good temperature tolerance. (Stark and Boyd, 1986)

Despite the lack of any clear knowledge of the precise mode of action, studies which have expressed the γ -subunit intracellularly from a GAL7, galactose inducible promoter region have demonstrated that it is this sub-unit which is responsible for the intracellular toxin activity. The alpha sub-unit is known to bind to cell walls and demonstrates a chitinase activity. (Butler *et al*, 1991)

1.8.3.2 Toxin Immunity

In order for toxin producing system to be viable in any context, the toxin producing strain must encode its own immunity mechanism. (Gunge *et al*, 1981) Within the k1 - k2 killer plasmid system several lines of evidence have demonstrated this to be the case. Evidence has suggested that the immunity gene for the killer toxin is encoded by k1. Further sub-cloning of killer plasmid DNA has demonstrated that k1 ORF3 appears to play a major role in immunity. The ORF3 region was sub-cloned onto a circular plasmid and introduced it into a $k2^+$ k1-free host. Transformants were found to have a significant but reduced level of toxin resistance than $k1^+$ $k2^+$ bearing cells: transformants had a 100% immunity to toxin concentrations up to 40ng/ml but only a 20-50% immunity to toxin at a concentration of 400 ng/ml, in contrast to wild-type, k1 containing cells which show a 100% resistance in toxin concentrations up to 400ng/ml. (Tokunaga *et al*, 1987) The analysis of this result demonstrated that wild-type k2 was also required for gene expression, thus k2 must either have a role in k1 ORF3 expression or produce a product which interacts with the ORF3 protein product.

1.9 Linear Plasmid Genetic Manipulations

Because of the linear and cytoplasmic nature of the k1 and k2 linear plasmids as well as the presence of a terminal protein moiety at the 5' ends of both strands, linear plasmids are not amenable to direct molecular manipulation techniques analogous to those used with circular plasmids. The key problems appears to lie not in the isolation of the linear plasmids, as several studies, in addition to this work have reported successful isolation of linear plasmids. (Stam *et al*, 1986, Reay, 1997) Only one study, so far, has reported the successful re-transformation of intact linear plasmid DNA and its subsequent replication *in vivo*. (Delouvencourt *et al*, 1983) Attempts to repeat this experiment in other labs have failed, thus suggesting that transformation is probably highly strain or condition dependent as well as being an extremely rare event. One of the causes of this inability to re-transform *K. lactis* with linear plasmids could be the loss of terminal protein molecules from the 5' ends. This leads us to the inevitable conclusion that, for example, a whole k1 plasmid, lacking terminal proteins could be transformed directly into a parental strain containing the k1 and k2 linear plasmids but would not be replicated and maintained if replication requires the presence of an existing terminal protein in order to initiate replication. There is also the problem that if the plasmids are transformed into a $k1^0$ $k2^0$ cell there will be no linear plasmid RNA polymerase present to allow transcription of linear plasmid genes.

Due to these difficulties with direct molecular manipulation, an alternative system has been devised to allow linear plasmid genetic manipulation. (Kamper *et al*, 1991, Shaffrath and Meacock, 1995) This technique involves the production of a recombinant linear plasmid fragment containing the desired recombinant region produced by cloning,

PCR and other molecular techniques on a circular *E. coli* based vector. This recombinant region must be flanked by linear plasmid DNA and a nutritional marker such as *TRP1* or *LEU2* fused to a linear plasmid UCS or promoter region allowing expression of these genes in the cytoplasm. These fragments must be arranged in such a way as to allow two targeted *in vivo* recombination events to occur thus resulting in the formation of a recombinant linear plasmid. This circular DNA molecule is then digested with appropriate restriction endonucleases to result in a linear recombination fragment. Following transformation of these linear fragments into a *K. lactis* strain, containing either the k1 and k2 plasmids or the k2 plasmid alone, recombinant linear plasmids are isolated following their selection on appropriate growth media. The result of this process is to produce cells harbouring both the wild-type and recombinant plasmids. Successive sub-culturing of the isolates allows the derivation of pure recombinant isolates, providing that the recombinant plasmid has not had an essential component deleted or deleteriously modified. Deleterious recombinants reach a state of equilibrium between the wild-type and recombinant plasmids, at an appropriate level to allow both the nutritional marker and essential gene product to be produced at the necessary levels. This genetic method, while cumbersome, is the best technique presently available for analysis of the linear killer plasmids and has been extensively used. (Schaffrath *et al*, 1992, Schaffrath and Meacock, 1995; Schaffrath and Meacock, 1996; Shaffrath *et al*, 1997) . Despite this technique's usefulness, it would require a great deal of perseverance to use for the fine analysis of genes or as a technique for the introduction of point mutations. The usefulness is also somewhat lessened by the availability of useful restriction endonuclease recognition sequences, although the use of PCR can overcome this.

One good example of the above genetic techniques being used for detailed analysis of a gene was the development by Schaffrath and Meacock, of a novel "gene shuffle"

method. k2 ORF5 was demonstrated to be an essential gene though a combination of disruption of the open reading frame and its transplacement onto k1, thus demonstrating its *trans*-active nature. This result was confirmed through immunochemical analysis carried out on the recombinant k2 ORF5 gene product which contained a C-Terminal c-myc epitope tag which could be detected using an anti-c-*myc* monoclonal antibody. (Schaffrath and Meacock, 1996)

method. k2 ORF5 was demonstrated to be an essential gene though a combination of disruption of the open reading frame and its transplacement onto k1, thus demonstrating its *trans*-active nature. This result was confirmed through immunochemical analysis carried out on the recombinant k2 ORF5 gene product which contained a C-Terminal c-myc epitope tag which could be detected using an anti-c-*myc* monoclonal antibody. (Schaffrath and Meacock, 1996)

1.10 Project Aims

The principal aims of this project were to explore the theory that the *Kluyveromyces lactis* killer plasmid terminal proteins are encoded within the cryptic 300 amino acids domains of the k1 and k2 DNA polymerase genes, k1ORF1 and k2ORF2. In order to approach this question, biochemical and genetic analysis was undertaken to relate the presence and properties of the amino-terminal domain of the DNA polymerase protein k2ORF2 to the known actions and localisation of the k2 terminal protein.

Chapter 2: Biochemical Analysis of the k2ORF2 DNA Polymerase

2.1 Introduction

2.1.1 Properties of the Killer Plasmid Terminal Proteins

My initial aim was to determine which of the fourteen K1/K2 encoded open reading frames or host genes encoded the k1 and k2 terminal proteins. Due to the fact that the killer plasmids can function in other strains and species of yeast, notably *Saccharomyces cerevisiae*, and the belief that the terminal proteins are essential for replication and plasmid stability, the idea that they would be host encoded seemed unlikely. (Gunge and Yamane, 1984; Stam *et al*, 1986; Gunge *et al*, 1995) If these proteins were host encoded then the genes would need to be very highly conserved across species; thus implying that these proteins would perform other key cellular functions.

2.1.2 Possible Locations of the Terminal Protein Genes

Because the two proteins show marked differences in both their electrophoretic mobility, size and their trypsin cleavage pattern, it is likely that they are either encoded by two separate open reading frames or are differentially processed from a common protein precursor. (Kikuchi *et al*, 1984; Stark *et al*, 1990) Computer analysis of the killer plasmid open reading frames has revealed none which would encode predicted polypeptides of precisely the correct sizes for the k1 terminal protein (28 kDa) and for the k2 (36 kDa) terminal protein. Also, if there were two genes which encoded the k1 and k2 terminal

proteins one would expect them to share a high degree of similarity if they had originated from the same ancestral gene. Amongst the k1 and k2 open reading frames however, there are no two, other than the two putative DNA polymerase, which share significant similarity.

Since all ORFs on plasmid k1 can be assigned functions, it seems unlikely that any could solely encode the k1 TP. Thus, if the two TPs are encoded by individual genes then those ORFs must be on k2.

This whole issue is further complicated by the discovery of another linear dsDNA plasmid which appears to exist alone in cells of the related yeast *S. kluyveri*, yet bears a striking similarity to k2. (Hishinuma and Hirai, 1991) This plasmid is called pSKL and is a 14.2 kb linear plasmid originally isolated from *Saccharomyces kluyveri* , a close relative of *Kluyveromyces lactis.* Plasmid pSKL DNA has a similar A + T content to k1 and k2 DNA, of 71.7%, and contains 10 large open reading frames which all appear to be transcribed. (Hishinuma and Hirai,1991) [see Figure 2.1.2.1] Furthermore, pSKL contains terminal inverted repeats and has covalently attached proteins, analogous to k1 and k2. (Hishinuma and Hirai, 1991) Interestingly, fifteen of the sixteen terminal nucleotides of pSKL are identical to those of k2, suggesting that the two plasmids may have similar methods of replication. (Hishinuma and Hirai,1991) Plasmids pSKL and k2 appear to be functionally identical. The open reading frames of the two plasmids are arranged in a similar orientation, are of similar size and are transcribed in the same direction. (Hishinuma and Hirai,1991) The predicted amino acid sequence for each pSKL open reading frame shares a high degree of similarity with that predicted for the corresponding k2 ORF product; for example, the putative k2 DNA Polymerase shares 61.7 % amino acid sequence identity to the putative pSKL DNA polymerase.

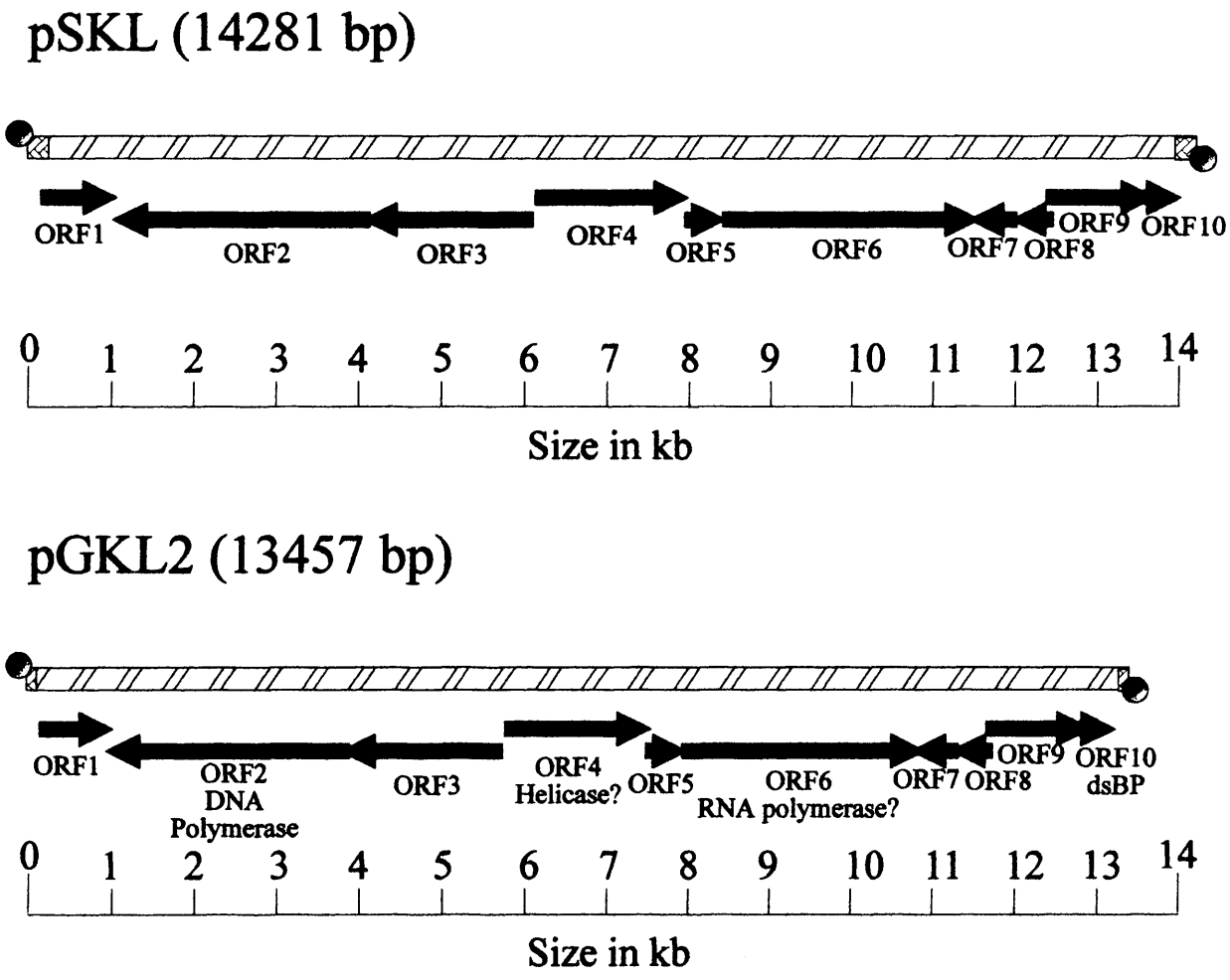


Figure 2.1.2.1: Genome Organisations of k2 and pSKL

Thus, the essentially analogous nature of the pSKL plasmid to k2 as well as the fact that all of the open reading frames appear to be highly conserved seems to argue against the k1 terminal protein being k2 encoded because of pSKL's solo nature in the cell.

An alternative theory is that the terminal proteins are encoded in the N-terminal regions of the k1 and k2 putative DNA polymerases. One of the bases of this theory comes from the composition of k1 and k2 predicted DNA polymerase polypeptide sequences. In addition to having strong amino acid identity with other B-Type viral DNA polymerases they are very similar to each other. (Stark *et al*, 1990; this study) This identity spans the central and carboxyl regions of polypeptide sequences but not the approximately 300 amino acids at the N-termini [Figure 2.1.2.2]. These non-homologous regions are capable of encoding polypeptides of approximately 28 kDa and 36 kDa for the k1 and k2 DNA polymerases, respectively, i.e the precise size of the k1 and k2 TPs.

This theory implies that each plasmid encodes its own DNA polymerase and explains the divergence of protein sequence at the amino terminal ends, where no DNA polymerase conserved sequences are located. Moreover, the N-terminal portions appear to have a characteristic amino acid composition for DNA binding proteins, with a high percentage of positively charged residues, such as lysine and arginine. See Figure 2.1.2.3 for the charge distribution of the polypeptide encoded by k2ORF2.

Therefore, a series of experiments were carried out to attempt to determine whether the amino-terminal region of the k2 DNA polymerase was the k2 terminal protein.

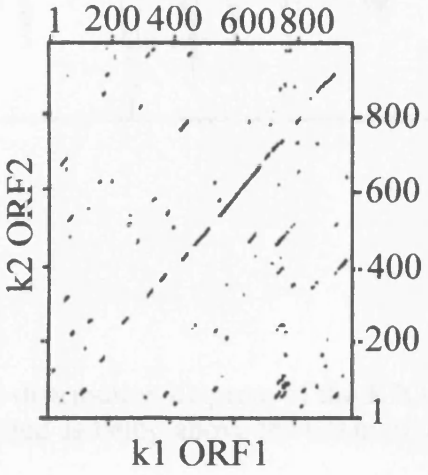


Figure 2.1.2.2: DotPlot comparison of the amino acid sequences of the k1 and k2 DNA Polymerases. [Window 30, Stringency 8]
(Stark et al, 1990)

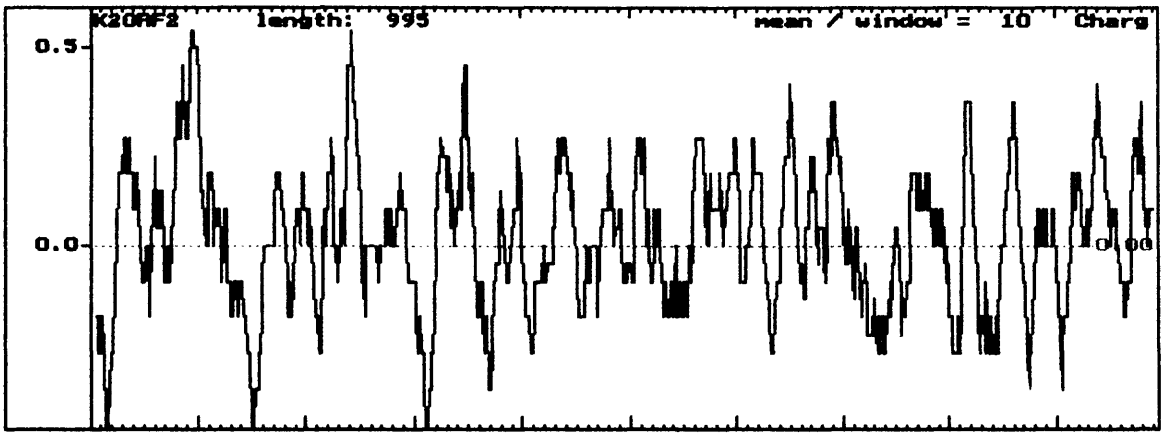


Figure 2.1.2.3: Charge distribution diagram of the K2ORF2 predicted amino acid sequence. Positive charge is depicted as being above the 0.0 mark and negative charges below.

2.2 Synthesis of a GST-k2ORF2 Overexpression Construct

The method we decided on, to investigate this theory, was an immunochemical approach. This was chosen, from several possibilities, mainly due to the limitations imposed on analysis by the killer plasmid system itself.

Due to the difficulty of expressing either the whole or fragments of the native killer plasmid genes in any standard expression system, a decision was made to use a fusion construct expression vector previously synthesised in the lab to attempt to over-express a fragment of the k2 putative DNA polymerase in *E. coli.* (C.A. Ambrose, PhD thesis; Romanos and Boyd, 1988) The construct was based on the expression vector pGEX-2T, in which the gene of interest is fused to a *Schistosoma japonicum* Glutathione S-Transferase (GST) construct driven by an IPTG-inducible *tac* promoter for expression in *E. coli.* (Smith and Johnson, 1988) [Figure 2.2.1 (Plasmid map of pGEX-2T)] In order to overcome the difficulties of expressing A + T rich killer plasmid open reading frames in *E. coli*, a 160 bp DNA fragment encoding the extreme N-terminus of the k2 putative DNA polymerase was re-synthesised utilising *E. coli* codon bias. (C.A. Ambrose, PhD thesis) Two 90-mer oligonucleotides sharing a 18bp overlap were annealed and PCR amplification resulted in a fragment encoding the N-terminal 54 amino acids of the K2 putative DNA polymerase, flanked by 5' *BamHI* and 3' *EcoRI* endonuclease recognition sequences. [Figure 2.2.2 (Construction of resynthesised fragment by PCR)] The fidelity of the PCR was confirmed by DNA sequencing. The fragment was purified, treated with *BamHI* and *EcoRI* endonucleases and sub-cloned into the pGEX-2T vector, resulting in an in frame fusion between the GST gene and the N-terminal fragment of the k2 DNA polymerase. (C.A. Ambrose, PhD thesis) The resulting plasmid was called pCA28. Expression studies were then performed. Although the 31 kDa fusion protein was

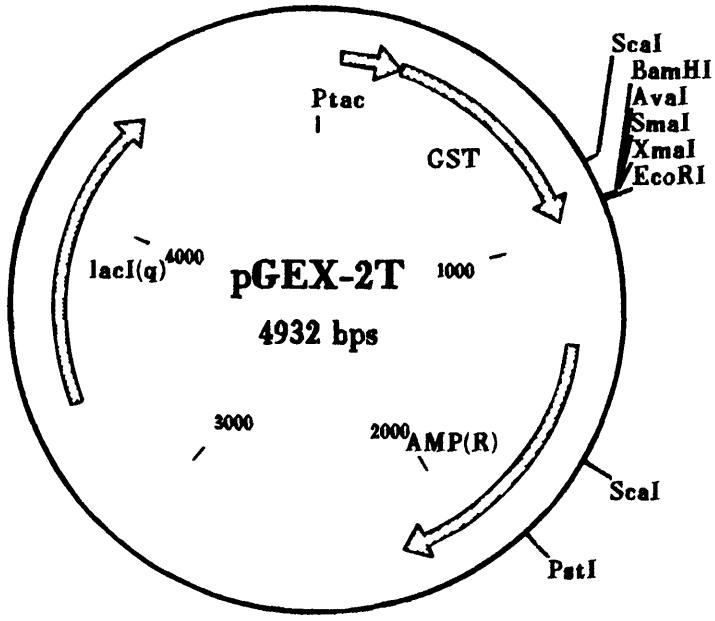


Figure 2.2.1: Plasmid map of pGEX-2T
This plasmid contains a *S.japonicum* Glutathione-S-transferase gene fused to a Ptac promoter with a lacI gene present to provide repression of GST gene expression. At the end of the GST gene is a multiple cloning site to allow in-frame fusion of the desired gene with the GST gene. The plasmid also contains an ampicillin resistance gene for the selection of *E. coli* transformants.

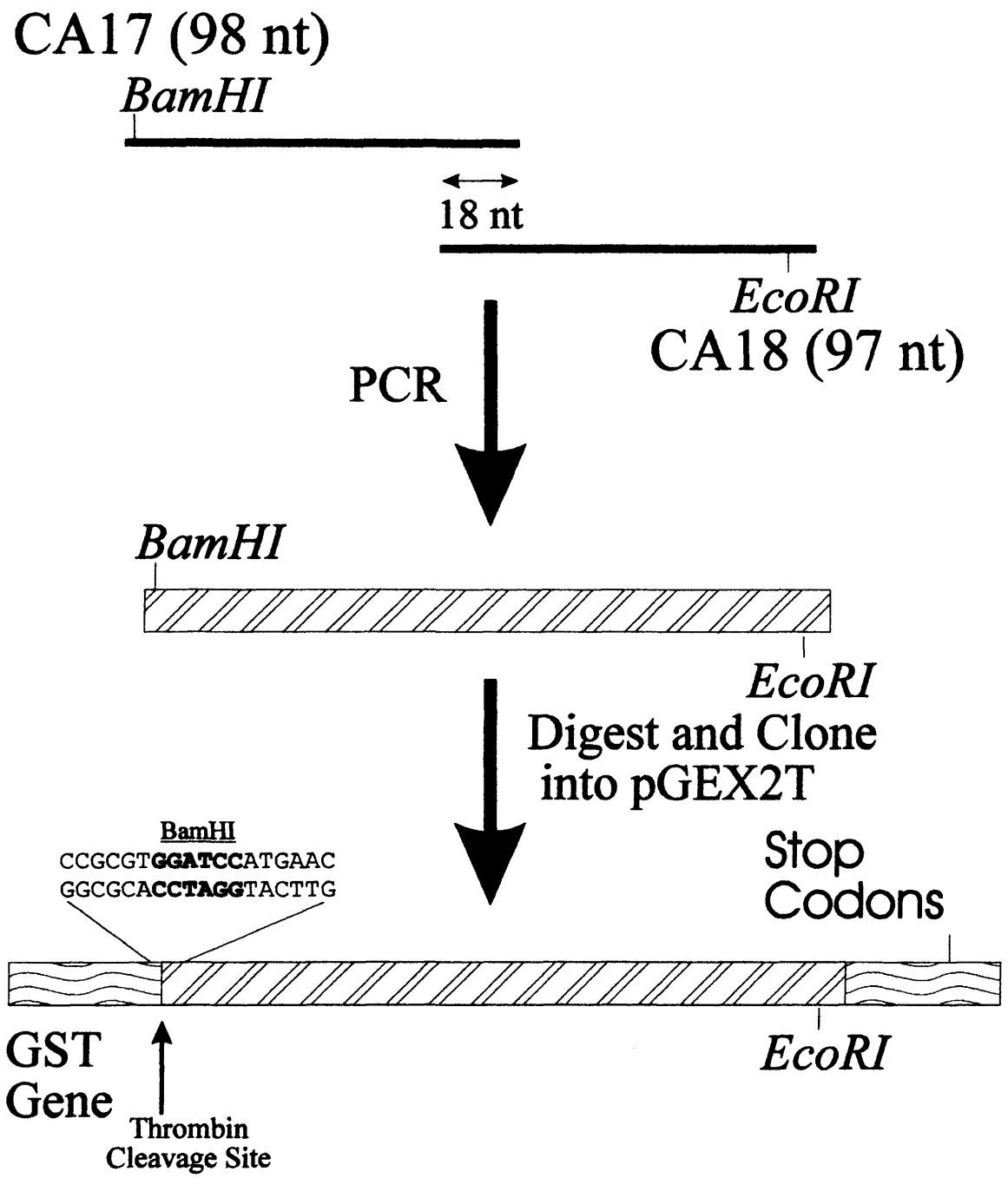


Figure 2.2.2: Construction of the GST-k2ORF2 fusion via PCR amplification of large oligonucleotides encoding the N-terminal 54 amino acids of k2ORF2.

expressed to a high percentage of total cell protein, initial attempts to purify this protein through the use of the intrinsic glutathione binding ability of the Glutathione S-Transferase failed. Because GST protein lacking the fusion domain was easily purifiable using this method, the conclusion drawn was that the resulting fusion protein must be insoluble. (C.A. Ambrose, PhD thesis)

2.2.1 Confirmation of the Identity and Authenticity of the GST-ORF2 Fusion

Upon starting my project, I resumed work on this construct with the goal of producing a protein of sufficient purity and in large enough quantities for antibody production. The first step was to confirm the plasmid's integrity.

2.2.1.1 Confirmation of the Construct by Restriction Digest

A frozen *E. coli* stock containing the plasmid was inoculated into growth media and plasmid DNA prepared. Purified pCA28 plasmid DNA was treated with *BamHI* and *EcoRI* restriction endonucleases and the resulting fragments analysed on a 0.7% agarose gel. The DNA was detected through the use of indirect fluorescence using Ethidium Bromide. Two fragments were detected, a 4 kb fragment and a 160 bp fragment, which agreed with the reported sizes of the vector and fusion fragment, respectively. (C.A. Ambrose, PhD thesis) [Figure 2.2.1.1.1]

2.2.1.2 Confirmation of the Construct by Protein Expression

Next, the plasmid was checked for its ability to express the fusion protein. Small scale IPTG induction experiments were performed and total cell protein (TCP) samples were analysed using SDS-PAGE. Following induction, bacteria containing the pGEX-2T vector produced a protein of approximately 27 kDa, while those containing the fusion

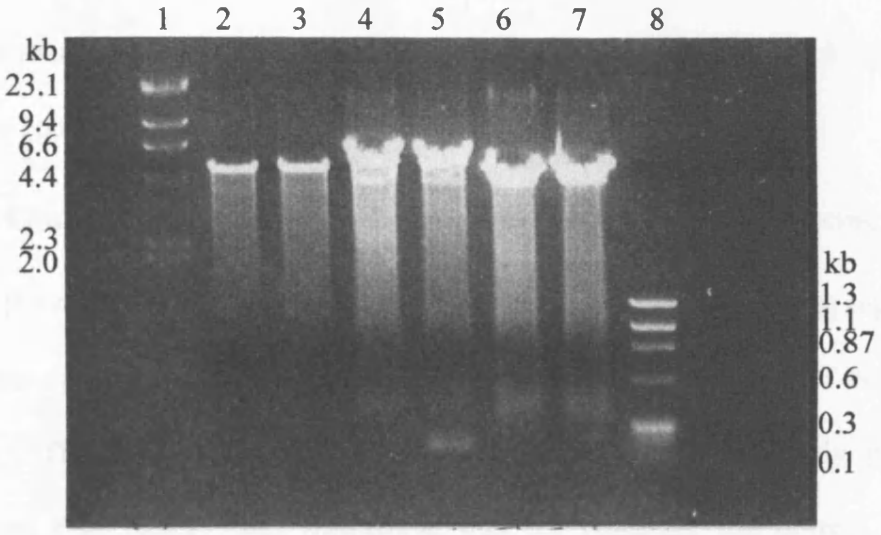


Figure 2.2.1.1.1 Agarose gel of plasmid pCA28 samples treated with restriction endonucleases. Lane 1: λ/*Hind*III molecular weight markers; Lane 2: pGEX-2T parental plasmid treated with *Eco*RI; Lane 3: pGEXT-2T parental plasmid treated with both *Bam*HI and *Eco*RI; Lane 4: pCA28 plasmid treated with *EcoRI;* Lane 5: pCA28 treated with *Eco*RI and *Bam*HI; Lane 6 and 7: pGEX-2T parental plasmid treated with *Eco*RI and *Eco*RI and *Bam*HI, respectively. Lane 8: ϕ/*Hae*III molecular weight markers

construct produced a protein of approximately 29 kDa. Both proteins were highly expressed and represented a significant proportion of total cellular protein [Figure 2.2.1.2.1]. Since the apparent molecular weight as detected via SDS-PAGE was 2 kDa smaller than the 31 kDa expected, DNA sequencing was performed to confirm that the plasmid construct was correct.

2.2.1.3 Confirmation of the GST-Fusion Construct By DNA Sequencing

An oligonucleotide identical to a portion of the upstream coding region of the GST gene was synthesised. Sequence data were then obtained for the GST-N-terminal fusion region. The sequence data obtained agreed with those previously published (C.A. Ambrose, PhD thesis), thus confirming that the construct was correct. The aberrant apparent molecular weight could be due to the highly charged nature of the protein affecting the correct resolution of its size by SDS-PAGE.

2.2.2 Purification of the Fusion Protein by Inclusion Body Isolation

Due to the reported insolubility of the GST-k2DPol fusion protein, in conjunction with its apparent abundance as judged through SDS-PAGE, it was deemed likely that this protein would form inclusion bodies in *E. coli*. Inclusion bodies are highly insoluble protein conglomerations formed through the interaction of the hydrophobic regions of partially folded proteins. They are relatively common when foreign proteins are overexpressed in *E. coli*. The highly insoluble nature of inclusion bodies can be used to purify them from other cellular proteins. Briefly, bacteria are lysed and the insoluble fraction of the lysate removed by high speed centrifugation. This insoluble fraction is then repeatedly washed with increasingly stronger detergents, to remove all but the most insoluble proteins, thus yielding a relatively pure inclusion body preparation. All the

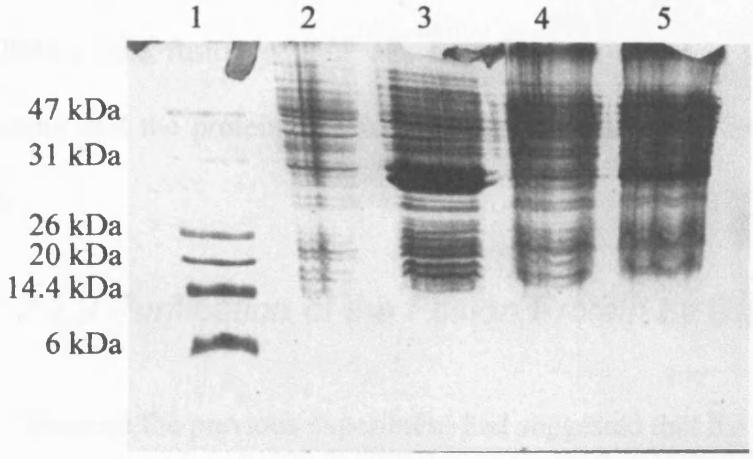


Figure 2.2.1.2.1: Total cell protein extracts of *E. coli* containing pGEX2T and pCA28 plasmids showing the expression of the GST and GST-k2ORF2 fusion proteins electrophoresed on a 10% SDS-PAGE gel, stained with Coomassie blue.
Lane 1: Low molecular weight protein standards, Lane 2: *E. coli* containing pGEX-2T before induction of GST expression; Lane 3: *E. coli* containing pGEX-2T following induction, Lane 4: *E. coli* containing pCA28 before induction of GST-fusion expression, Lane 5: *E. coli* containing pCA28 following induction of GST-fusion expression.

fractions are then assayed for the desired protein to judge at which point the protein is separated from the insoluble mass, or to determine protein loss through this treatment if the protein is almost entirely in inclusion bodies.[see Figure 2.2.2.1]

Following induction of protein synthesis through the addition IPTG, an inclusion body preparation was performed. When this preparation was analysed using SDS-PAGE, very little of the fusion protein was found to be present in the insoluble fraction, thus suggesting that the protein was soluble. Repetitions of this experiment yielded identical results.

2.2.3 Purification of the Fusion Protein by Glutathione Affinity

Because the previous experiment had suggested that the fusion protein was soluble, it seemed likely that it could be purified by using the intrinsic affinity of the GST for glutathione sepharose. In order to purify large quantities of the GST-k2ORF2 fusion protein 5ml of L-Broth containing 50 μg/ml ampicillin were inoculated with the host *E. coli* strain containing the fusion plasmid and incubated overnight. The next day, samples were diluted 1:100 into fresh L-Broth supplemented with 50 μg/ml Ampicillin and grown for two hours. Expression of the fusion protein was then induced by the addition of IPTG to a final concentration of 1mM and the culture was grown for a further 3 hours. Cells were then harvested by low speed centrifugation and disrupted by sonication. Cell debris was removed from the supernatant via high speed centrifugation and the supernatant was poured through a column containing Glutathione sepharose beads. The column was washed with 10 volumes of column buffer and fusion protein eluted with column buffer containing 10 mM glutathione. Following purification, the purified protein samples were analysed using SDS-PAGE and appeared to be virtually free of other protein species.

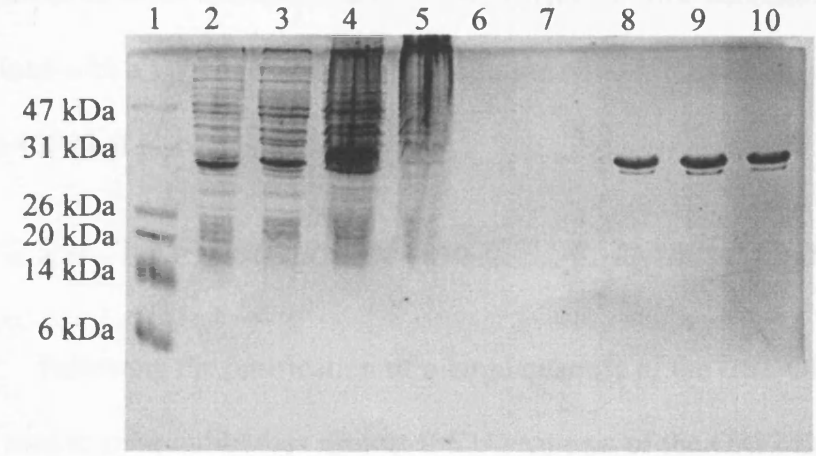


Figure 2.2.2.1: Inclusion body preparation from induced pGEX2T and pCA28 containing strains *E. coli* strains electrophoresed on 10% SDS-PAGE gel, stained with Coomassie blue. Lane 1: low molecular weight standard; Lane 2: Total cell protein extract from pGEX2T strain, uninduced; Lane 3: Total cell protein extract from pCA28 strain, uninduced; Lane 4: Insoluble fraction from pGEX-2T strain; Lane 5: Insoluble fraction from pCA28; Lane 6: pGEX-2T diluted soluble lysate fraction. Lane 7: pCA28 diluted soluble lysate fraction; Lane 8 and 10: protein prepared from pGEX-2T via GST affinity; Lane 9: protein prepared from pCA28 via GST affinity.

[Figure 2.2.3.1(Purification of GST-Fusion Proteins)] The 29 kDa GST and 31kDa fusion, however appeared to have very similar migration patterns on the 15% SDS-PAGE gel. This is due to the polyacrylamide concentration not showing separation characteristics within this molecular weight range to separate the 29 kDa and 31 kDa GST and GST-Fusion proteins. Later SDS-PAGE analysis used an acrylamide concentration of 10% which showed better resolving properties within this range. Finally, a protein assay was performed in order to determine preparation yield; a final concentration of 453 μg/ml was obtained with a total of 5ml of protein solution obtained, resulting in a total yield of 2.2mg from 400ml of *E. coli.*

2.2.4 The Production of Anti-GST-N-terminal ORF2p antibodies

Following the purification of a large quantity of the GST-ORF2 fusion protein, this was used to raise antibodies against the N-terminus of the ORF2 DNA-polymerase. After study and consideration of the various animal systems available for the production of polyclonal-antibodies, chickens were judged to be the most suitable for this series of experiments, due to their superior resilience, reduced antigen requirement and high antibody production. (Gassman *et al*, 1990) The requirement for polyclonal antibodies was based on our need for a highly sensitive antibody due to the very low level of the terminal protein present in the cell (2 per linear plasmid molecule). A course of two injections was performed. Two weeks after the course was completed, eggs were obtained and IgY, the chief constituent of the chicken antibody response, was purified using a differential PEG_{6000} precipitation method described by Gassman *et al* (1990). Pre-immune IgY was also purified from eggs taken from the two chickens prior to the first antigen injection.

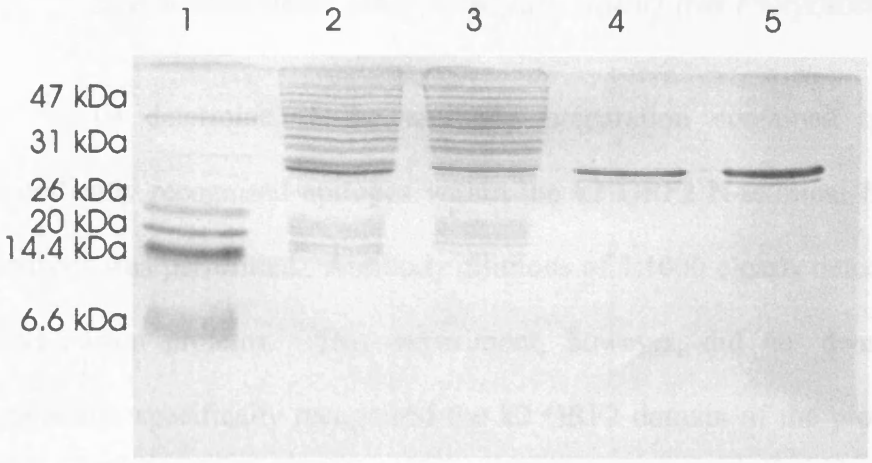


Figure 2.2.3.1: Purification of GST-Fusion proteins as shown by SDS-PAGE analysis. Electrophoresis performed on 15% SDS-PAGE gel and silver stained. .Lane 1: low molecular weight protein standards. Lane 2:Total cell protein pGEX-2T containing *E. coli* induced. Lane 3: Total cell protein pCA28 containing *E. coli* induced. Lane 4: Glutathione purified protein from pGEX-2T containing *E. coli*. Lane 5: Glutathione purified protein from pCA28 containing *E. coli*

2.2.5 *Western Blot Analysis using the Polyclonal Antibody*

To determine if the antibody preparation contained IgY molecules which specifically recognised epitopes within the k2 ORF2 N-terminal fragment, Western-blot analysis was performed. Antibody dilutions of 1:1000 clearly detected both the GST and GST-Fusion proteins. This experiment, however, did not demonstrate whether any antibodies specifically recognised the k2 ORF2 domain of the protein since most of the antigen injected was GST protein.

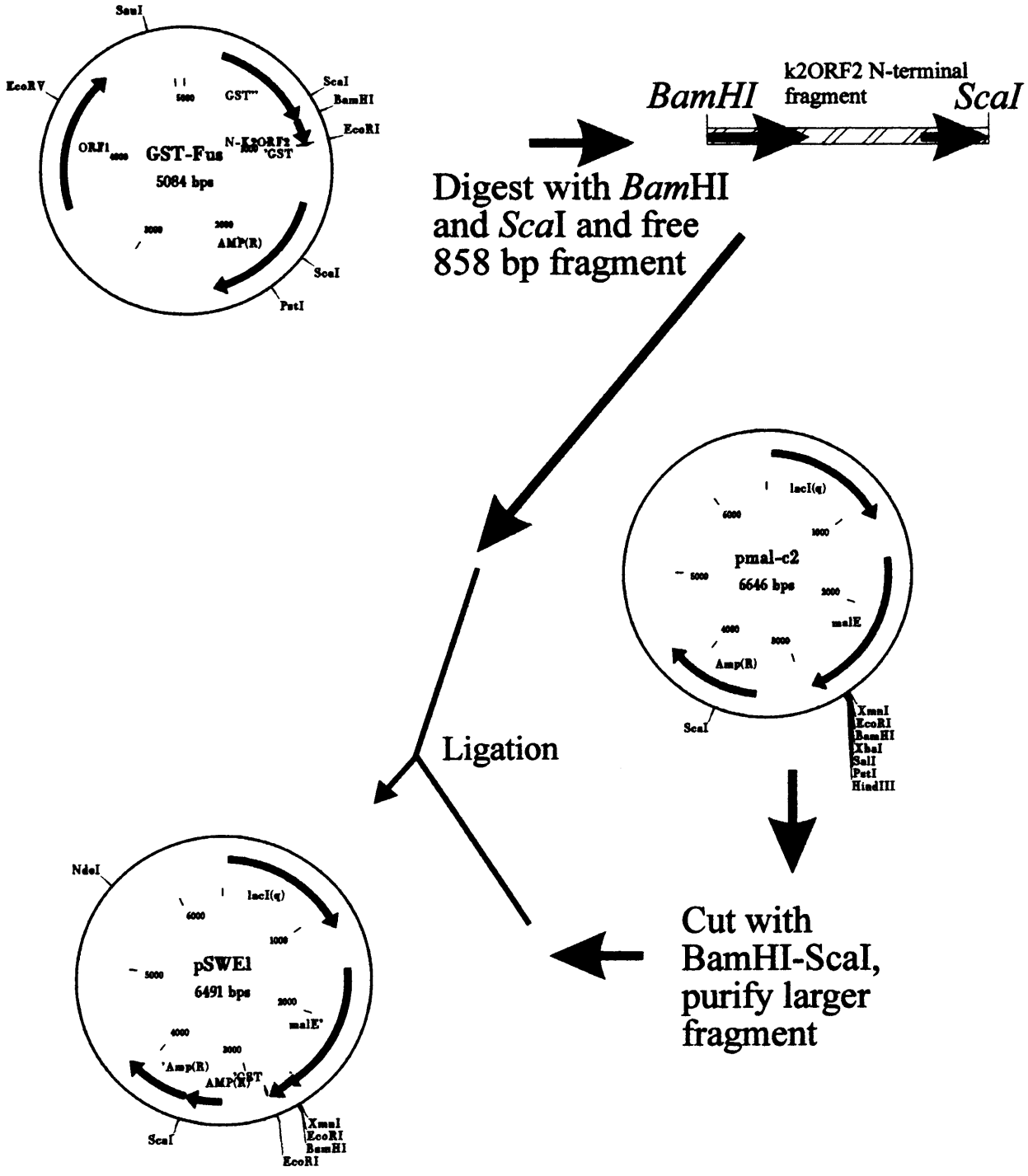
In order to demonstrate that specific anti-k2ORF2 antibodies were present, an attempt was made to separate the GST and ORF2 peptides through the use of the Thrombin protease consensus target sequence, located between the GST and ORF2 domains. When these experiments were carried out, however, although Thrombin clearly cleaved the fusion protein, only the 27 kDa product, which corresponds to the GST carrier protein could be detected in Western analysis. Silver staining of Thrombin treated fusion protein clearly showed that no detectable amount of the 5 kDa fragment, which corresponds to the re-synthesised ORF2 fragment, was produced. The two possible explanations for this were that the small peptide is being degraded by proteases during its incubation, or that Thrombin, in addition to its normal cleavage site, was cleaving a site within the fragment thus resulting in one or more fragments, of a size below the level of detection for conventional SDS-PAGE. The optimal thrombin cleavage sites are: A-B-Pro-Arg||-X-Y where A and B are hydrophobic amino acids and X and Y are non-acidic amino acids and Gly-Arg||-Gly. (Chang, 1985) There were, however, no obvious matches for these two patterns within the k2ORF2 fusion region.

2.3 Synthesis of a Maltose Binding Protein (MBP)-ORF2 Fusion Protein

Due to the difficulty encountered in obtaining separate GST and k2ORF2 amino terminal protein fragments using thrombin protease cleavage, a DNA construct capable of producing a second fusion protein was designed utilising the Maltose Binding Protein fusion vector pMAL-C2. This was constructed on the assumption that the anti GST-k2ORF2 antibody preparation would have no antibody molecules which would specifically recognise the *E. coli* maltose binding protein. Thus if an MBP-ORF2 fusion protein was detected by Western analysis using the antibody obtained from the previous inoculation of GST-ORF2 this interaction would, most likely be the result of a specific interaction with the ORF2 peptide part of the fusion protein.

In order to express this new protein a new *E. coli* expression vector was constructed. The first step in the construction of this vector was the treatment of the pCA28-GST fusion construct with *BamHI* and *ScaI* endonucleases. This resulted in the liberation of two fragments, 868 bp and 4.1kb in size. The 868 bp fragment contained the entire k2 ORF2 fragment, 14 bp of the pGEX2T vector and a 305 bp fragment of the β-lactamase gene; this k2 ORF2 fragment was then purified from a 0.7% agarose gel. The pMAL-c2, *E. coli* Maltose Binding Protein vector was also treated with the *BamHI* and *ScaI* endonucleases, resulting in the production of a 5.6kb fragment containing the vector portions of the plasmid and an unwanted, 1kb insert fragment. The 5.6kb fragment was purified and ligated to the 868 bp ORF2 containing fragment described earlier, thereby recreating the ampicillin resistance gene thus resulting in a new MBP-ORF2 fusion construct, pSWE1 [see Figure 2.3.1]. This construct was then confirmed using restriction analysis.

Figure 2.3.1: Construction of the MBP-k2ORF2 expression plasmid pSWE1



2.3.1 SDS-PAGE Characterisation of MBP-N-terminus k2 ORF2 Transformants

Following selective growth on medium containing 50 μg/ml ampicillin, resistant *E. coli* transformants were picked and sub-cultured in 2ml test-tubes. Once the culture reached an optical density (A_{600}) of 0.4, synthesis of the fusion protein was induced by the addition of IPTG to a final concentration of 0.1mM. All of the 12 transformants picked, appeared to produce a 49 kDa fusion protein, to approximately 50% of total cell protein, as judged by SDS-PAGE analysis. The fusion protein was somewhat smaller than the original MBP fusion, because the MBP protein replaced in the original vector was fused to a fragment of the lacZ gene, resulting in a MBP-β-galactosidase of 51kDa. The new fusion construct disrupted the original MBP-*lacZ* fusion.

2.3.2. Western Analysis of a MBP-ORF2 Fusion Transformant

Maltose binding protein and MBP-ORF2 fusion proteins were both purified using amylose affinity column chromatography in a method analogous to the column affinity purification of the GST and GST-fusion proteins. The main difference between the two methods is the use of an amylose bead resin and elution conditions containing 10mM maltose. Yield of fusion protein from this procedure was much lower than with the GST-process, with an MBP-fusion yield of approximately 200 μg/ml in 5 ml or a total of 1mg of protein from 400ml. Proteins were eluted from the column and their purity assayed using SDS-PAGE analysis followed by band visualisation via silver staining. [Figure 2.3.2.1(SDS-PAGE analysis of Purified MBP-Fusion protein)] The MBP-ORF2 sample had few co-purifying contaminants, but the MBP sample had several contaminating

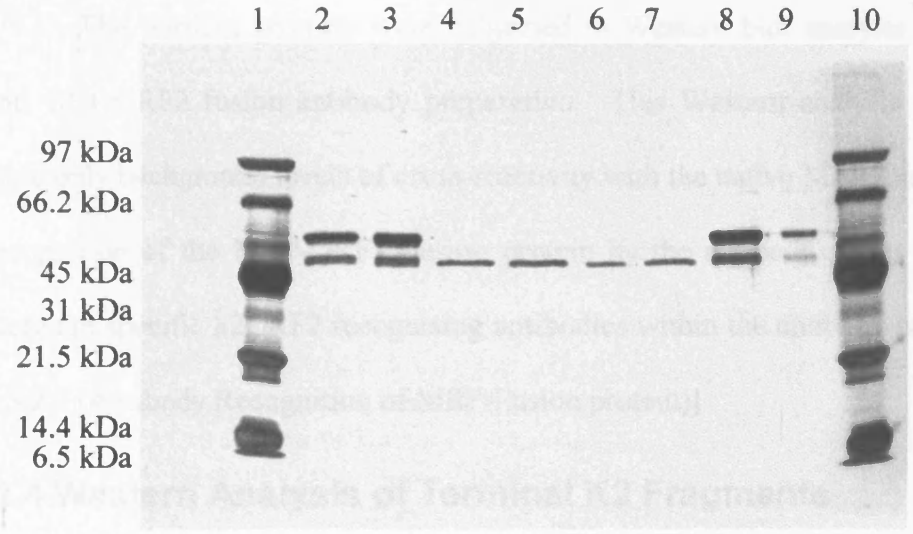


Figure 2.3.2.1: Purified Maltose binding protein and maltose binding protein k2ORF2 fusion protein analysed on a 10% SDS-PAGE gel and silver stained.. Lane 1 and 10: Protein low molecular weight standards; Lane 2 and 9: Purified Maltose binding protein. Lane 3 and 8: Purified maltose binding protein following Factor Xa treatment. Lane 4 and 7: Purified MBP fusion protein Lane 5 and 6: Purified MBP fusion protein following factor Xa treatment.

proteins which constituted less than 1% of total cell protein. In order to further purify the MBP sample, approximately 10 μg of the purified MBP sample were electrophoresed on a 10% SDS-PAGE gel, blotted onto a PVDF (Polyvinyl Diethyl Flouride) membrane and the ~49 kDa band purified using a PVDF band purification method (see Materials and Methods). The resulting sample had no contaminants as visualised by silver stain analysis.

The purified proteins were subjected to Western blot analysis using the chicken anti GST-ORF2 fusion antibody preparation. This Western analysis showed that there were only background levels of cross-reactivity with the native MBP, but there was a clear recognition of the MBP-ORF2 Fusion protein by the antibody. This demonstrated that there are specific k2ORF2 recognising antibodies within the antibody preparation. [Figure 2.3.2.2 (Antibody Recognition of MBP-Fusion protein)]

2.4 Western Analysis of Terminal K2 Fragments

Following the determination that the antibody preparation did contain specific, anti-k2orf2p antibodies, the next step was to investigate the presence of a cross-reacting protein at the termini of the k2 plasmid. Such a finding would provide good evidence that the k2 terminal proteins are derived from the primary k2ORF2 gene product. Several different methods were considered and attempted for the isolation of intact k2 linear plasmids and one method was finally determined to be the most effective based on the percentage of k2 which still had attached terminal protein, as determined by agarose mobility studies. This was an adaptation of a method previously developed in this laboratory for the preparation of linear plasmid DNA. It is described in detail in the Materials and Methods chapter; briefly it involves the gentle lysis of linear plasmid containing *Kluyveromyces lactis* cells, followed by fractionation of the cytosolic

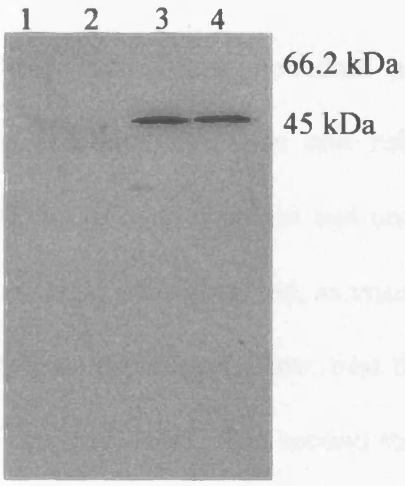


Figure 2.3.2.2 Western analysis of MBP fusion protein following electrophoresis on a 10% SDS-PAGE gel and Western transfer onto a nitrocellulose membrane visualised using the LUMIN8 chemi-luminescence kit. The antibody used was the anti-GST-Fusion protein antibody present at concentration of 1:10,000.
Lane 1: Purified MBP protein
Lane 2: MBP protein following Factor Xa treatment
Lane 3: Purified MBP-Fusion protein
Lane 4: MBP-fusion protein following Factor Xa treatment
(Factor Xa did not cleave a significant proportion of the MBP-fusion protein)

components via Ammonium Sulphate precipitation resulting in the enrichment of the sample for the linear plasmids against other nucleic acids and proteins. Further purification can be performed by subjecting the resultant sample to size fractionation on a Sepharose S-1000 fine column.

During development of the purification procedure parameters such as culture growth conditions, precipitating salt concentrations and cell wall digestion were all optimised. The procedure was found to be very robust and under most conditions similar quantities and purities of the k2 plasmid were obtained, as visualised by SDS-PAGE. Two different strains were used for this isolation procedure: first the wild type $k1^+$ $k2^+$ strain *IFO1267* and second the $k1^0$ $k2^+$ strain *SD802*. The second strain was expected to be the more useful for this series of experiments due to the need to obtain pure k2 plasmid DNA, but in my hands, the yield and purity of k2 plasmid was never as high as that obtained from the wild-type, IFO1267 strain. Therefore the IFO1267 strain was used preferentially.

Experiments were then performed, attempting to detect the k2 terminal protein directly with the anti-GST-k2ORF2 antibody. A starting amount of 5 μg of purified linear plasmid DNA was electrophoresed on a 4% polyacrylamide gels and then transferred to a nitrocellulose membrane. Western blot analysis using a primary antibody concentration of 1:1000, however showed no detectable signal. Twenty repetitions of this experiment using antibody concentrations of 1:100 to 1:10,000 and plasmid DNA concentrations of up to 50 μg were carried out. However, direct detection of the terminal protein, was ever achieved and when low antibody dilutions of 1:100 - 1:200 were used, very high non-specific background signal became apparent. In order to attempt to increase the binding of DNA to the blotting substrate, Western and Southern blotting onto Nylon membranes was also attempted. Nylon has a significantly increased binding affinity for DNA compared

with Nitrocellulose. This result yielded disappointing results as a very high level of background signal was detected during antibody analysis of the resulting filter. Experiments were also performed using Agarose supplemented with 2% v/v SDS and by performing DNAse treatment prior to SDS-PAGE, but these were also unsuccessful.

At about this time a paper was published by Takeda *et al* detailing the isolation of a large quantity of the k2 terminal protein and its subsequent amino acid sequencing. The limited sequence data they obtained, due, they claim to the high amount of degradation of the k2 terminal protein following its liberation from the k2 DNA, demonstrated that the terminal protein was derived from the N-terminal region of the k2 ORF2 predicted polypeptide. (Takeda *et al*, 1996)

2.5 Discussion

The results published by Takeda *et al* (1996) are not totally conclusive because they predict the start of the terminal protein to be well inside the N-terminal region of the K2 ORF2 polypeptide at amino acid 228. This would suggest that, based on its known size, the terminal protein would extend through several conserved DNA polymerase domains. [see Figure 2.5.1] They do, however, provide the first clear evidence that the k2 terminal protein is derived from the amino terminal region of k2ORF2.

Suggestions as to why my experiments failed are strongly hinted at by the problems encountered by the group which carried out the sequencing experiments. In order for them to obtain the 10 mg of k2 TP they finally obtained, and were able to sequence, they had to start with 1kg of *Kluyveromyces lactis* cells. This acceptable, but low yield of terminal protein would suggest that there is a high level of loss of linear plasmid during the purification procedure. I calculate that this yield represents about 10 - 20% of the total amount of k2 DNA. A large part of the loss of linear plasmid during purification could be due to the inherent "stickiness" observed by both our group and other investigators. (Reay, 1997) This characteristic is also evident during electrophoresis as k2 linear plasmids with attached terminal proteins do not migrate through an agarose gel matrix in the absence of high salt, such as 150mM-200mM NaCl, or ionic detergents, such as 0.1% SDS. Following treatment with a non-specific protease, such as protease K, this "stickiness" is no longer seen. This effect complicates further manipulation of the plasmid DNA. This "stickiness" coupled with the difficulty in isolating sufficient amounts of the protein species means that despite the avidity of the antibody or the sensitivity of the detection method a very large amount of k2 plasmid would be needed for any meaningful detection,

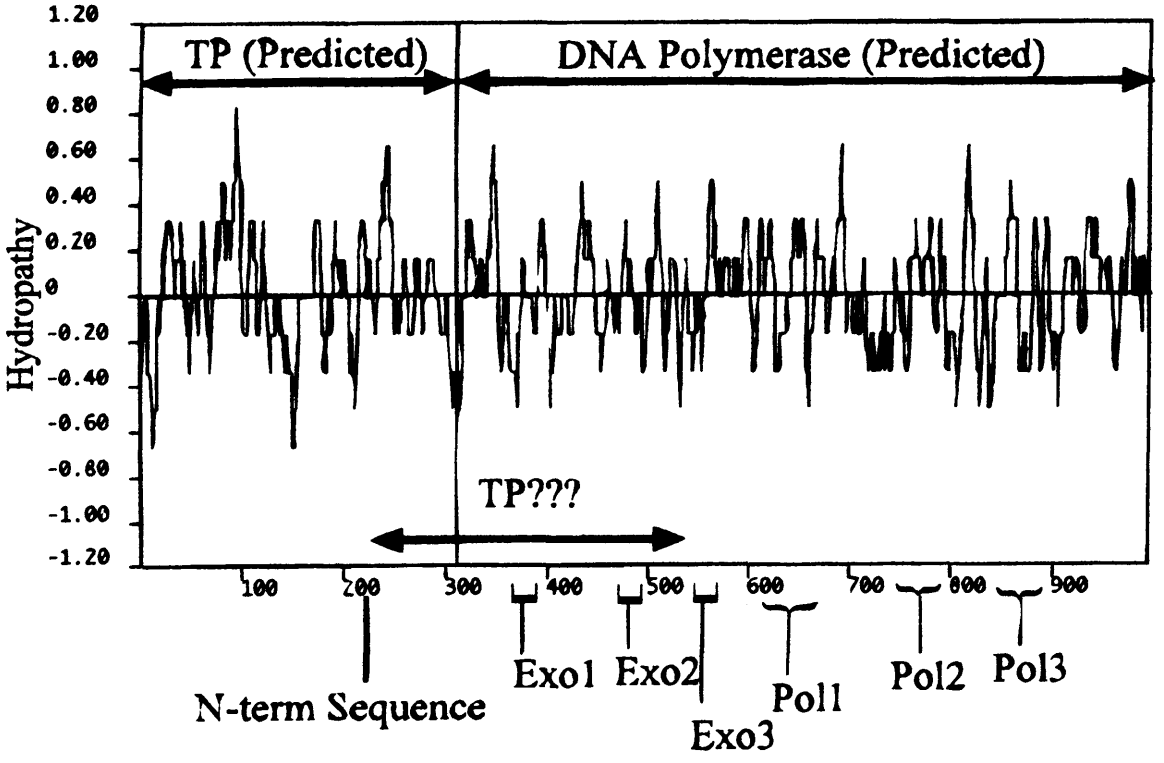


Figure 2.5.1: Comparison of the predicted domain structure of the k2ORF2 peptide product. TP(Predicted) shows the predicted terminal protein based on its known size. DNA Polymerase (Predicted) illustrates the predicted DNA polymerase domain containing the conserved EXO (exonuclease) and POL (polymerase) domains. TP??? shows the extent of the terminal protein based on N-terminal sequence data obtained by Takeda *et al* for the terminal protein. Note the overlap of the predicted terminal protein and the two exonuclease domains.

especially because the polyclonal antibody raised only recognises a small region of the putative terminal protein.

Chapter 3: Genetic Analysis of the k1ORF1 DNA Polymerase

3.1 Introduction

Sequence analysis has implicated k1ORF1 and k2ORF2 open reading frames as encoding DNA polymerases. (Jung *et al*, 1987; Tommasino *et al*, 1988) Earlier disruption analysis carried out by R. Schaffrath in this laboratory has demonstrated that the k2ORF2 putative DNA polymerase is indeed a gene essential for the maintenance of k2. (Shaffrath and Meacock, 1995) This suggests, together with previous evidence demonstrating the independence of k2 maintenance from that of k1, that the two DNA polymerase serve identical functions for the two different plasmids. In order to confirm this theory, therefore, it was necessary to demonstrate that the putative k1ORF1 DNA polymerase was essential for k1 maintenance, but not needed for k2 replication.

3.2 Construction of a k1ORF1 Disruption Cassette in *Escherichia coli*

In order to carry out the disruption of the k1ORF1 open reading frame it was necessary to construct a k1ORF1 disruption cassette bearing a selectable marker, in this case the LEU2 leucine biosynthesis gene, for positive selection of the mutant plasmid. All genetic analysis on linear plasmids, as noted before, must follow a certain pattern. The cycle of a gene disruption involves four key steps:

- Construction of a disruption cassette containing the disrupted gene flanked both upstream and downstream with linear plasmid DNA to target correct integration. A selectable marker is also contained within the DNA fragment to be inserted, driven by a linear plasmid UCS promoter element. [Figure 3.2.1]
- Liberation of the DNA fragment carrying the disruption cassette and transformation into the target host *K. lactis* cell [Figure 3.3.1].
- Selection of transformants and molecular characterisation to ensure correct integration of the disruption cassette into the target linear plasmid.
- Successive selective sub-culturing of transformants and genetic analysis to determine the phenotypic effect of the gene disruption.

Since the initial recombination event only occurs in a single plasmid molecule, the long term successive sub-culturing under selective conditions is needed to drive the plasmid population to homogeneity or at least to an equilibrium situation. This work intensive approach to gene disruption is necessitated due to the high copy number,

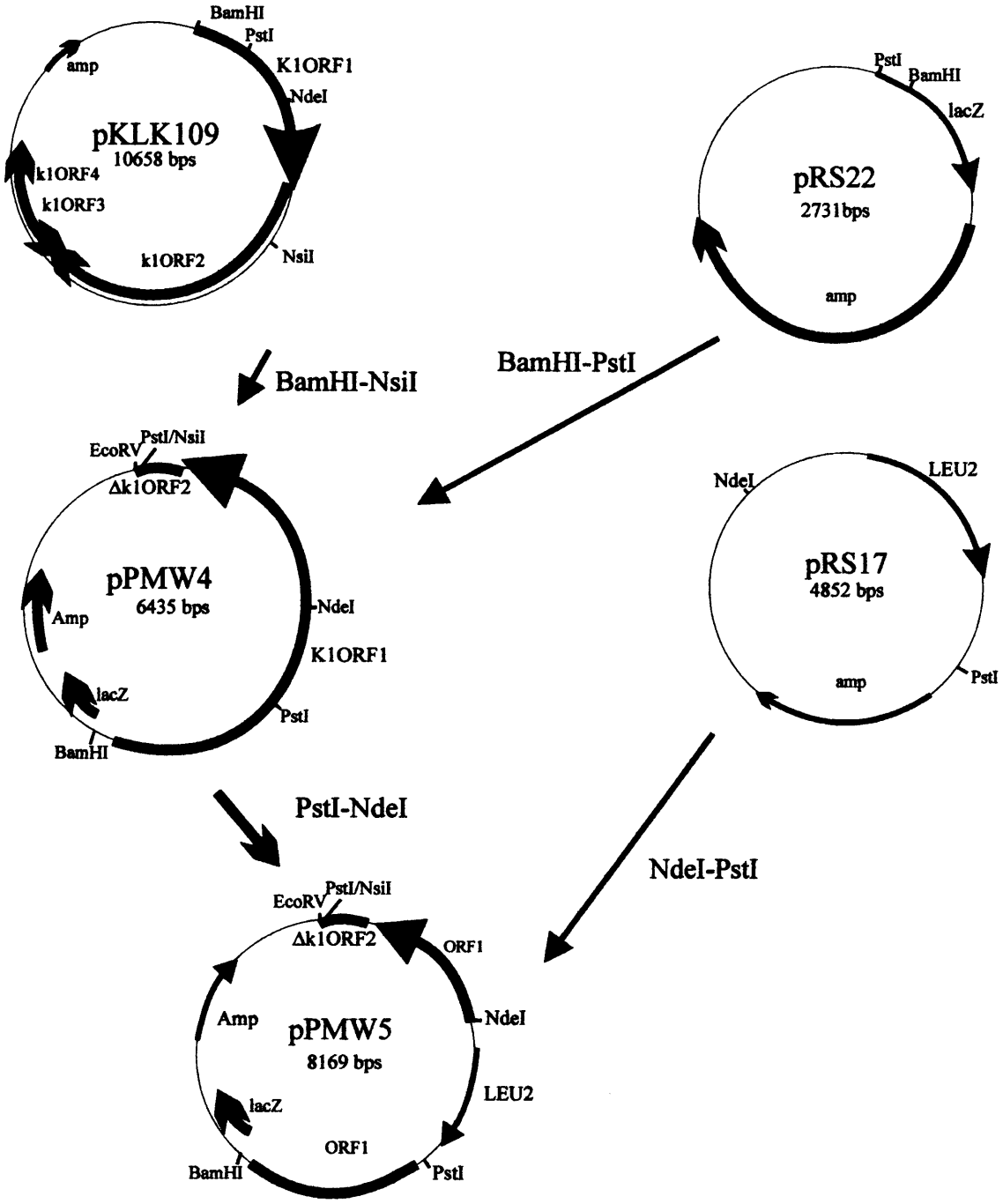


Figure 3.2.1: Construction of the pPMW5 k1ORF1 disruption vector. The k1ORF1 gene was sub-cloned from the parental pKLK109 plasmid into the pUC based vector pRS22 to form pPMW4.
The *Saccharomyces cerevisiae* LEU2 gene which is fused to a killer plasmid UCS promoter element was then transferred from its parental pRS17 vector into pPMW5 thereby causing a large central deletion of the k1ORF1 gene.

approximately 100 copies per cell, of the linear plasmids within a linear plasmid bearing yeast cell.

The first step in the construction of the k1ORF1 disruption cassette was the sub-cloning of a fragment, containing all of the k1ORF1 into a suitable vector for further manipulation [see Figure 3.2.1]. This was done by treating 10 μg of pKLK109 plasmid DNA with *Bam*HI and *Nsi*I restriction enzymes to yield a 4428 bp fragment which contained both the k1ORF1 gene and approximately 1 kb of k1ORF2. Next 10 μg of pRS22 plasmid DNA was treated with *Bam*HI and *Pst*I restriction enzymes. Plasmid pRS22 is a pUC-based vector with the *Nde*I and *Hind*III restriction enzyme target sites deleted. Successful ligation together of the compatible *Pst*I and *Nsi*I "sticky ends" results in the disruption of both recognition sites. The band purified k1ORF1 fragment and the digested pRS22 vector were then ligated together using standard techniques, using a insert to vector ratio of 4:1. Following an overnight ligation reaction the mixture was transformed into competent *E. coli* and trasformants were selected for by plating on Luria agar media supplemented with 50 μg /ml ampicillin. Following overnight growth, 24 colonies were picked and grown overnight in 2ml of Luria broth supplemented with 50 μg /ml ampicillin. The next day small scale plasmid preparations were performed using 1ml of each culture. The plasmid DNA was then subjected to restriction enzyme treatment using *Eco*RV and *Bam*HI. The samples were then electrophoresed on a 0.7% agarose gel and stained with Ethidium bromide. The expected result was that for a successful ligation two bands should be present following restriction, a 2.7 kb pRS22 vector band and a 3.8 kb k1ORF1 band. Four of the samples showed the correct banding pattern and one of these was chosen for further plasmid manipulation. The resultant plasmid was designated pPMW4. The next step was the deletion of a central 651bp fragment of ORF1 by the

digestion of the pPMW4 plasmid with *Pst*I and *Nde*I. The LEU2 for insertion into the *Pst*I-*Nde*I gap was obtained from a *PstI-NdeI* digestion of the pRS17 vector, followed by the preparatory electrophoresis and purification of the UCS-LEU2 containing fragment. The two fragments were then ligated together yielding the final disruption cassette contained within a pUC based vector in *E. coli.* This plasmid was called pPMW5 [Figure 3.2.2]. The plasmid structure was confirmed by restriction analysis [see Figure 3.2.3].

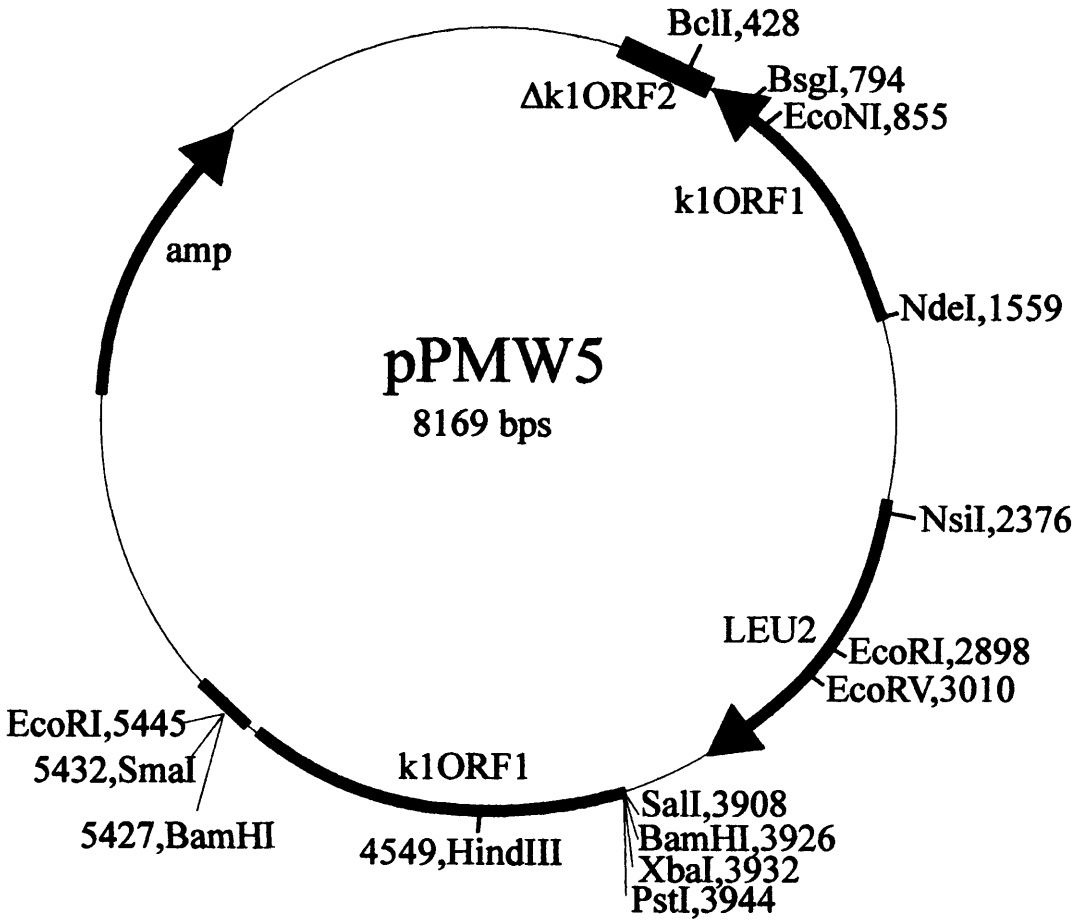


Figure 3.2.2: The k1ORF1 disruption plasmid pPMW5

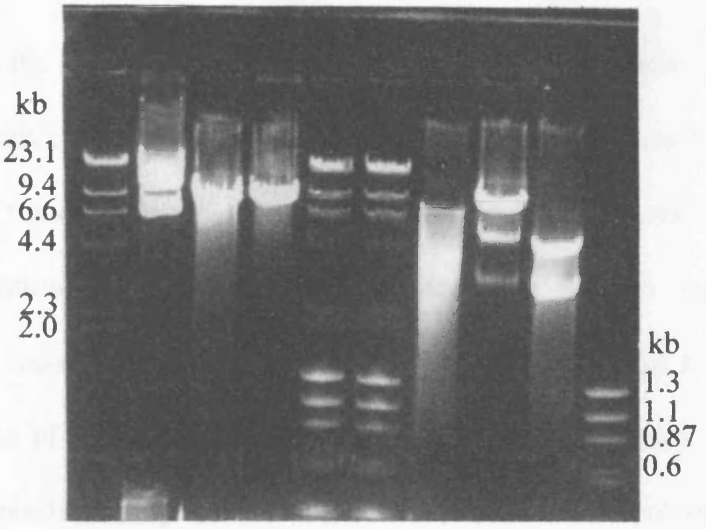


Figure 3.2.3: Restriction enzyme analysis of the k1 disruption vector pPMW5
Lane 1: λ/*Hind*III DNA molecular weight standards Lane 2: untreated plasmid Lane 3: *Bcl*I treated plasmid Lane 4: *Xma*I treated plasmid Lane 5 and 6: λ/*Hind*III molecular weight markers and ϕX174/*Hae*III molecular weight markers. Lane 7: *Sma*I treated plasmid Lane 8: *Bcl*I-*Xma*I treated plasmid (digestion did not run to completion) Lane 9: *Bcl*I-*Sma*I

3.3 Disruption of the K1ORF1 gene in *Kluyveromyces lactis*

Once the k1ORF1 disruption cassette had been constructed in *E. coli*, the next step was to introduce it into a host *K. lactis* strain which contained both k1 and k2. The liberation of the cassette and its transformation into the host strain allowed it to take part in recombination with the k1 linear plasmid. The cassette was constructed so that homologous recombination would target it to integrate into the k1 plasmid replacing the central region of the k1ORF1 gene with a UCS-LEU2 construct which can produce the LEU2 gene product using the intrinsic linear plasmid RNA polymerase. The disruption cassette lay within a 4.1 kb *Hind*III - *Bcl*I fragment. [Figure 3.3.1]

3.3.1 Production of Δk1ORF1 Disruptant Strains

A large scale *Hind*III - *Bcl*I restriction digest was performed using 50 μg of pPMW5 plasmid DNA. Following treatment, the digestion mixture was added directly to two 700 μl aliquots of competent *K. lactis* and a transformation performed. The *K. lactis* strain used was AWZ-137 as it contains both k1 and k2 and the useful *leu2* and *trp1* auxotrophic markers. Following transformation, leucine prototrophic transformants were recovered on synthetic define medium (SD) plates supplemented with tryptophan, but lacking leucine. This should select for integration and expression of the Δ-k1ORF1:UCS:*leu2* cassette. The transformation mix was also plated onto tryptophan and leucine-containing medium as a positive control, and medium containing neither tryptophan nor leucine as a negative control. A third control of cells which had not been exposed to any vector DNA was also used to ascertain the frequency of spontaneous Leu^+ revertants.

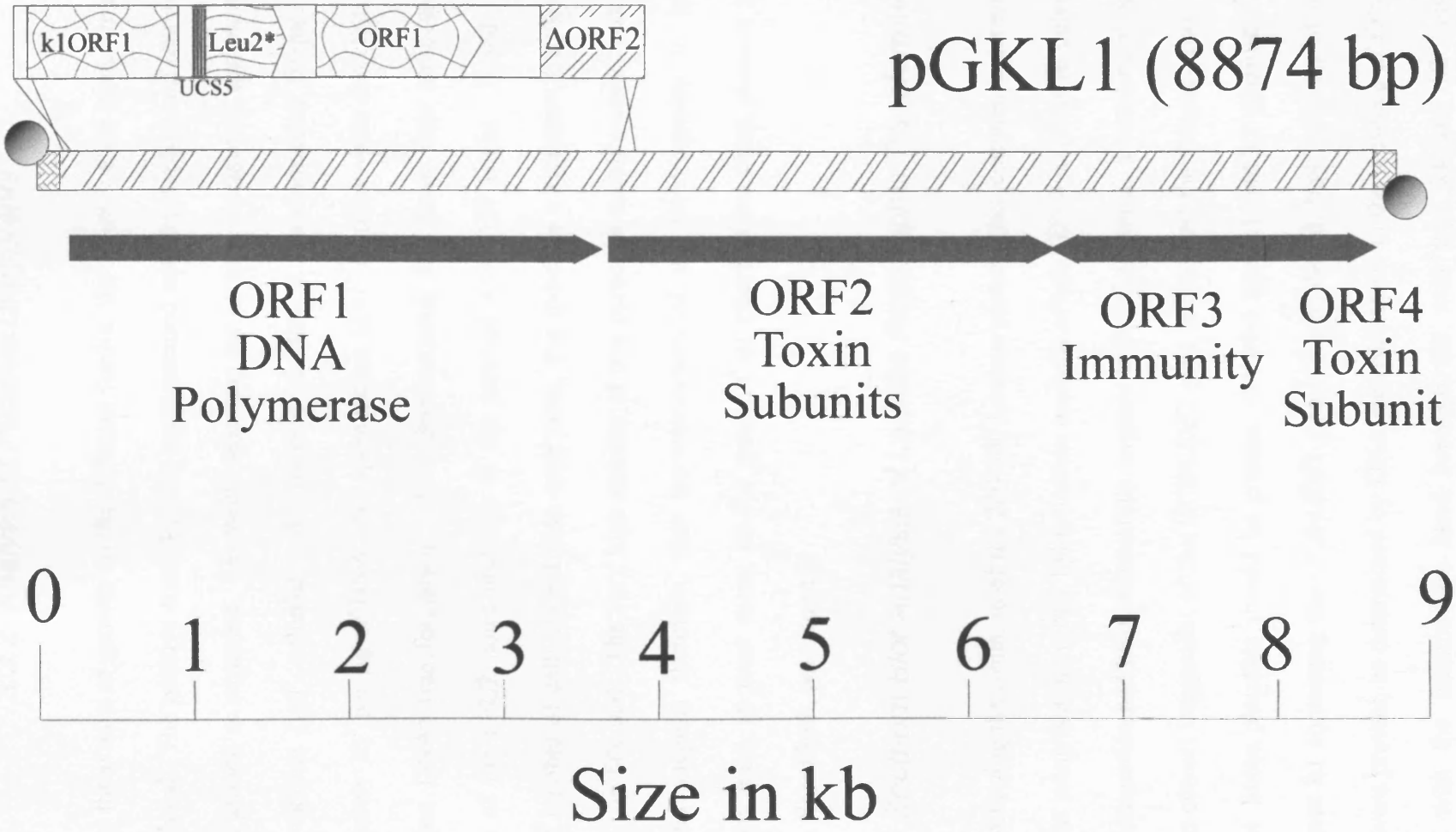


Figure 3.3.1: Plan for disruption of the k1ORF1 gene using a disruption cassette liberated from pPMW5

3.3.2 Analysis of Putative Disruptants

Following three days of growth on the different media, the results of the transformations were analysed. The positive control plates supplemented with tryptophan and leucine had confluent growth as expected. On media selecting for cassette integration approximately 200-300 distinct Leu^+ colonies were counted yielding a transformation frequency of approximately 10^2 per μg of DNA. No spontaneous LEU^+ colonies were detected in the control (no DNA) transformation. This demonstrated that there was no detectable reversion of the LEU2 point mutation in the parental AWZ-137 strain. Based on the estimated number of cells spread onto each plate, this indicates a reversion frequency of less than 1 in 600,000. The very high stability of this particular mutation has been noted before and suggests therefore, that the conversion of the transformants to leucine prototrophy must, in most cases, be the result of an effect of the input plasmid vector DNA. (R. Schaffrath, pers. comm)

3.3.3 Southern Blot Analysis of Leucine Prototrophic Transformants

Following the isolation of the 200-300 Leucine prototrophic colonies, a screen was carried out using Southern blot hybridisation analysis of total DNA extracts to determine whether integration had been successfully targeted to the k1 plasmid. The expectation was that if the correct integration of the UCS-LEU2 gene had caused the leucine prototrophy then three linear plasmids should be present, the wild type k1 and k2 plasmids and a recombinant k1 containing the UCS-LEU2 gene. As predicted, both versions of the k1 plasmid were present, as determined by DNA electrophoresis in the presence of Ethidium bromide, with the intermediate band between the wild type k1 and k2 bands at

approximately 9.5kb in size, as expected. Southern analysis using a *S. cerevisiae* LEU2 probe revealed that the LEU2 gene was contained within the recombinant plasmid, as expected [Figure 3.3.3.1]. Southern hybridisation was performed as detailed in the Materials and Methods chapter. For the initial screen, a ^{32}P-labelled *Saccharomyces cerevisiae* LEU2 probe was used to detect marker integration. This screen provided some interesting results. In most cases, the LEU2 marker had not properly integrated into the k1 plasmid. Hybridisation was detected to both low molecular weight, and extremely high molecular weight, DNA species. The low molecular weight structures detected could be circularised linear plasmid DNA as linear plasmid DNA is capable of forming circular structures and replicating via the use of fortuitous ARS regions contained within it. The high molecular weight fragments detected could be chromosomal or integrated LEU2 sequences. In two isolates, however hybridisation was detected to correctly sized, recombinant k1 plasmids.

3.3.4 Sub-culturing of a Δk1ORF1 Strain

The next stage in the analysis of the isolated disruptant strains was to attempt displacement of the wild-type k1 plasmid from the strain containing the k1ORF1 deletion by continuous rounds of growth and selection for the recombinant Δ-k1ORF1-UCS:LEU2 k1 plasmid. This is a key part of the analysis which tests whether the disrupted gene is essential for plasmid maintenance; the logic being that a wild type k1 plasmid can support the growth of a recombinant k1 plasmid that lacks essential genes provided the gene product can function in *trans* (i.e. is diffusable). In this situation, however the loss of the

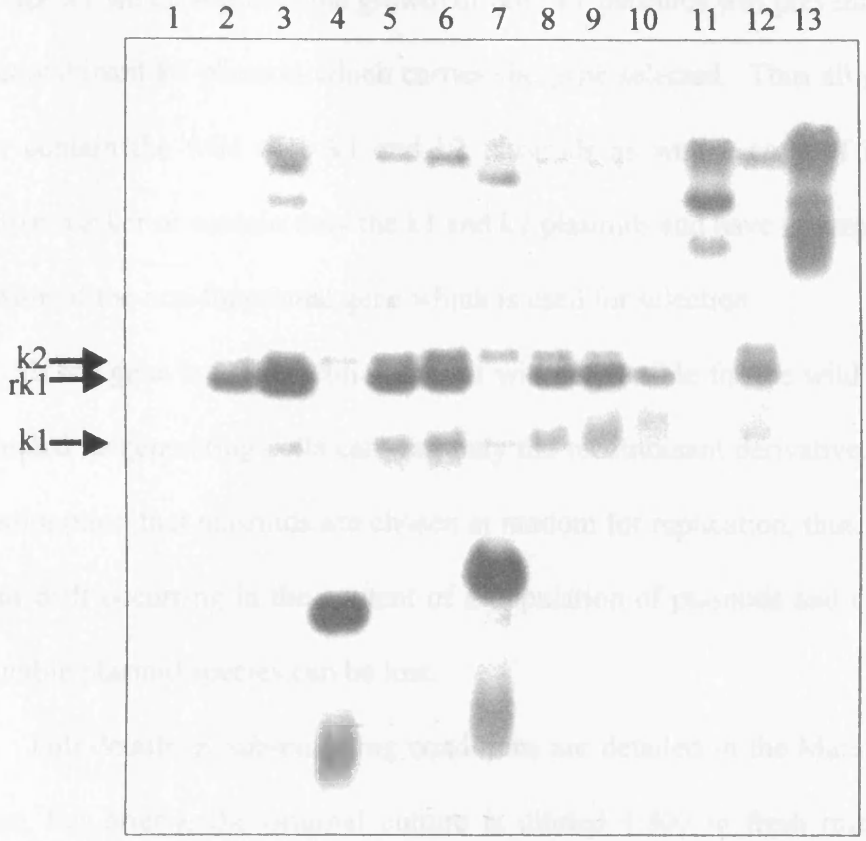


Figure 3.3.3.1: Southern blot analysis of total DNA samples treated with RNase and protease of LEU$^+$ transformants of AWZ-137 obtained using the pPMW5 disruption cassette. Samples electrophoresed on 0.7% agarose gel supplemented with 0.1% SDS and Southern blotted onto a nylon membrane and probed with *Saccharomyces cerevisiae* LEU2 probe.
k1, k2 and rk1 show the locations of the k1, k2 and recombinant k1 plasmid on the autoradiograph. This autoradiograph was over-exposed in order to allow the visualisation of the wild type linear plasmid species. There are no easily visible k1 and k2 plasmids in Lane 1 because there was significantly less growth of this strain. Longer exposures, however, did reveal the presence of the k1 and k2 plasmids at the locations shown.
Lane 1: AWZ-137 parental strain
Lane 2-12: LEU$^+$ isolates from pPMW5 disruption cassette transformation .Transformant samples in lanes 2,3,5,6,8,9 and 12 show a recombinant rk1 plasmid.
Lane 13: untreated pPMW5 plasmid DNA as a positive control

wild type k1 which supports the growth of both k1 plasmids will prevent the replication of the recombinant k1 plasmid which carries the gene selected. Thus all growing cells will either contain the wild type k1 and k2 plasmids as well a copy of k1 containing the selective marker or contain only the k1 and k2 plasmids and have undergone a spontaneous reversion of the non-functional gene which is used for selection.

If the gene is non-essential, then it will be possible for the wild type species to be eliminated so generating cells carrying only the recombinant derivative. This is based on the assumption that plasmids are chosen at random for replication, thus one gets a gradual bias or drift occurring in the content of a population of plasmids and eventually the non-favourable plasmid species can be lost.

Full details of sub-culturing conditions are detailed in the Materials and Methods section, but briefly, the original culture is diluted 1:800 in fresh minimal media (SD) supplemented with glucose and tryptophan but lacking leucine, thus selecting for maintenance of the ORF1Δ Leu^+ recombinant k1 plasmid. This process of subsequent sub-culturing was continued for approximately 250 generations. Following each round of sub-culturing linear plasmid DNA was purified and kept for later analysis by agarose gel electrophoresis and hybridisation. The results of this analysis demonstrated that one strain had lost its recombinant plasmid, but the other still had both the recombinant and wild-type k1 plasmids present at levels apparently very similar to those at the start of the selective growth regimen. [Figure 3.3.4.1 (Final southern analysis of Sub-culturing). Since neither strain had lost its wild type plasmid in favour of the recombinant under selective pressure, this strongly suggests that the putative k1 DNA polymerase is an essential gene product. This finding correlates with data obtained by Shaffrath and Meacock on k2ORF2 which also encodes a putative DNA polymerase. (Shaffrath and Meacock, 1995)

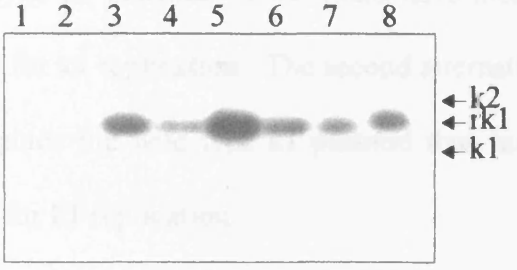


Figure 3.3.4.1: Southern blot analysis of linear plasmid samples taken from LEU$^+$ pPMW5 disruption cassette transformants. Samples were electrophoresed on a 0.7% agarose gel and blotted onto a nylon membrane for analysis. The probe used was *Saccharomyces cerevisiae* LEU2. The clear rk1 band can be seen continuing throughout the rounds of sub-culturing. The presence and locations of the k1 and k2 bands were established from other, over-exposed, autoradiographs and Ethidium bromide visualisation.
Lane 1: AWZ-137 parental strain
Lanes 3-8: purified linear plasmid samples isolated from rk1 containing cultures following 0, 10, 20, 30, 40, 50 and 60 generations, respectively.

3.4 Discussion

What was sought in this work was a clear indication as to whether k1ORF1 is an essential gene. Two results could have been seen: the first was that a disrupted k1ORF1 containing k1 plasmid bearing the wild type *S. cerevisiae* LEU2 gene could replicate successfully within the yeast cell and, through random selection for replication in some cases displace the wild-type k1 plasmid. This would have meant that the k1ORF1 gene product was not required for k1 replication. The second alternative was that the disruptant derivative could not displace the wild type k1 plasmid thus indicating that the k1ORF1 gene product is essential for k1 replication.

This series of experiments have suggested that the k1 ORF1 putative DNA polymerase is indeed essential for k1 maintenance and correlates with data obtained for the k2 putative DNA polymerase disruption. Also they supported the idea that two DNA polymerases are required for the killer plasmid system. The question then immediately arises, why two? Why has selective pressure not reduced the number of DNA polymerases to one? As was mentioned in the last chapter, the N-terminal region appears to hold the key to the differentiation of the different DNA polymerases as it is this region which is the most divergent. Further comparison between the highly homologous k2 and pSKL plasmids which have very similar N-terminal ends as well as a highly similar ITR structure suggests that the ITR sequences could be the reason for this difference as terminal proteins and ITR sequences must co-evolve. Further experiments which could be carried out to clarify this relationship would be to swap the ITR regions of the k1 and k2 plasmids and repeat this experiment although this would be very difficult to do with the current state of linear plasmid technology. In order to do this for k1 one would need to flank the k1 DNA with k2 terminal DNA and attempt to use homologous recombination and selection to

form a recombinant plasmid *in vivo*. The second alternative would be to attempt to construct an entire k1 plasmid with k2 ITR regions and an internal selectable marker and transform it into a k2 only auxotrophic strain. Successful transformants could be selected via plating onto selective medium.

If the ITR regions are truly responsible for this difference then the wild-type k1 plasmid should be able to be displaced by the recombinant LEU2 containing k1 plasmid.

Chapter 4: Production and Characterisation of a Hybrid k1ORF1-k2ORF2 DNA Polymerase

4.1 Introduction

The products of k1ORF1 and k2ORF2 are homologous and appear to be related to the type B (viral) DNA polymerases. The only significant differences between them occur in the amino-terminal 1/3 of the proteins, whereas the central and carboxyl-terminal regions, which contain the conserved exonuclease and polymerase motifs, are closely related. Experimental evidence has suggested that each of the DNA polymerases cannot compensate for a severe defect in the other. The question that inevitably arises is where does the plasmid specificity lie within each of the two DNA polymerases. (Schaffrath *et al*, 1995) Presumably, this specificity must lie within the approximately 300 amino acid, non-conserved, amino-terminal region. Interestingly the amino terminal region of k2ORF2 has also been implicated as encoding the terminal protein. (Takeda *et al*, 1996; this study) Therefore, it was decided to try to carry out a "domain-swap" experiment to ask if the exonuclease and polymerase domains were interchangeable.

In order to determine the significance of the amino terminal amino acid sequence divergence, comparisons were done between the k2 and pSKL plasmid. The plasmid, pSKL is closely related to k2 and has both a very high degree of terminal inverted repeat sequence homology as well as a very high level of DNA polymerase similarity throughout the entire predicted protein sequence. [Figure 4.1.1: (Comparison of k2 versus k1 and pSKL DNA polymerase)]

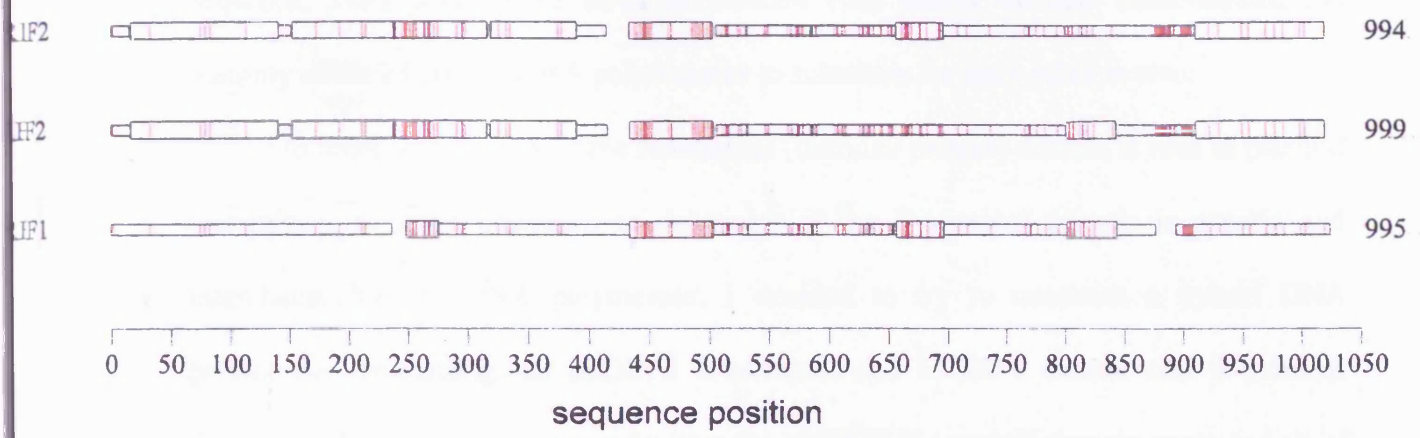


Figure 4.1.1: Predicted polypeptide sequence comparison between k1ORF1, k2ORF2 and pSKL ORF2 putative DNA polymerases.

Evidence from deletion studies of both the k2 DNA polymerase (Schaffrath and Meacock, 1995) and the k1 DNA polymerase (this study) has also demonstrated the inability of the k1 and k2 DNA polymerases to substitute for each other *in vivo*.

In order to determine if the N-terminal (terminal protein) domain is vital to plasmid recognition, or more precisely, to determine if the C-terminal domain is generic and interchangeable for DNA polymerase, I decided to try to construct a hybrid DNA polymerase, containing the k2ORF2 N-terminal and k1ORF1 central and C-terminal domains. The region encompassed within the k2ORF2 N-terminal domain contained all of the non-conserved regions as well as a small region of the conserved central DNA polymerase region.

4.2 Construction of the k2ORF2-K1ORF1 DNA Pol Hybrid Gene

The pMS100 plasmid contains the k1 "left" inverted terminal repeat, as well as a small (300 bp) fragment of the k1ORF1 gene, immediately flanked by a LEU2 nutritional marker driven by a k1ORF2 UCS. This plasmid was chosen due to its advantageous restriction sites and because it already contained k1 sequences, as well as a selectable marker. See Figure 4.2.1 Construction of pSWE3.

Next, plasmid DNA from the plasmid pKLK109, which contains a fragment of the k1 plasmid starting from the "left-hand" inverted terminal repeat and continuing to the start of the k1ORF4 gene, was isolated from the dam$^-$ *E. coli* strain, JM110. The Dam methylase of *E. coli* performs methylation of adenosine residues within a GATC tetramer. This methylation can inhibit the activity of some restriction enzymes which contain this partial sequence in their recognition site, such as *BclI* and *XbaI*.

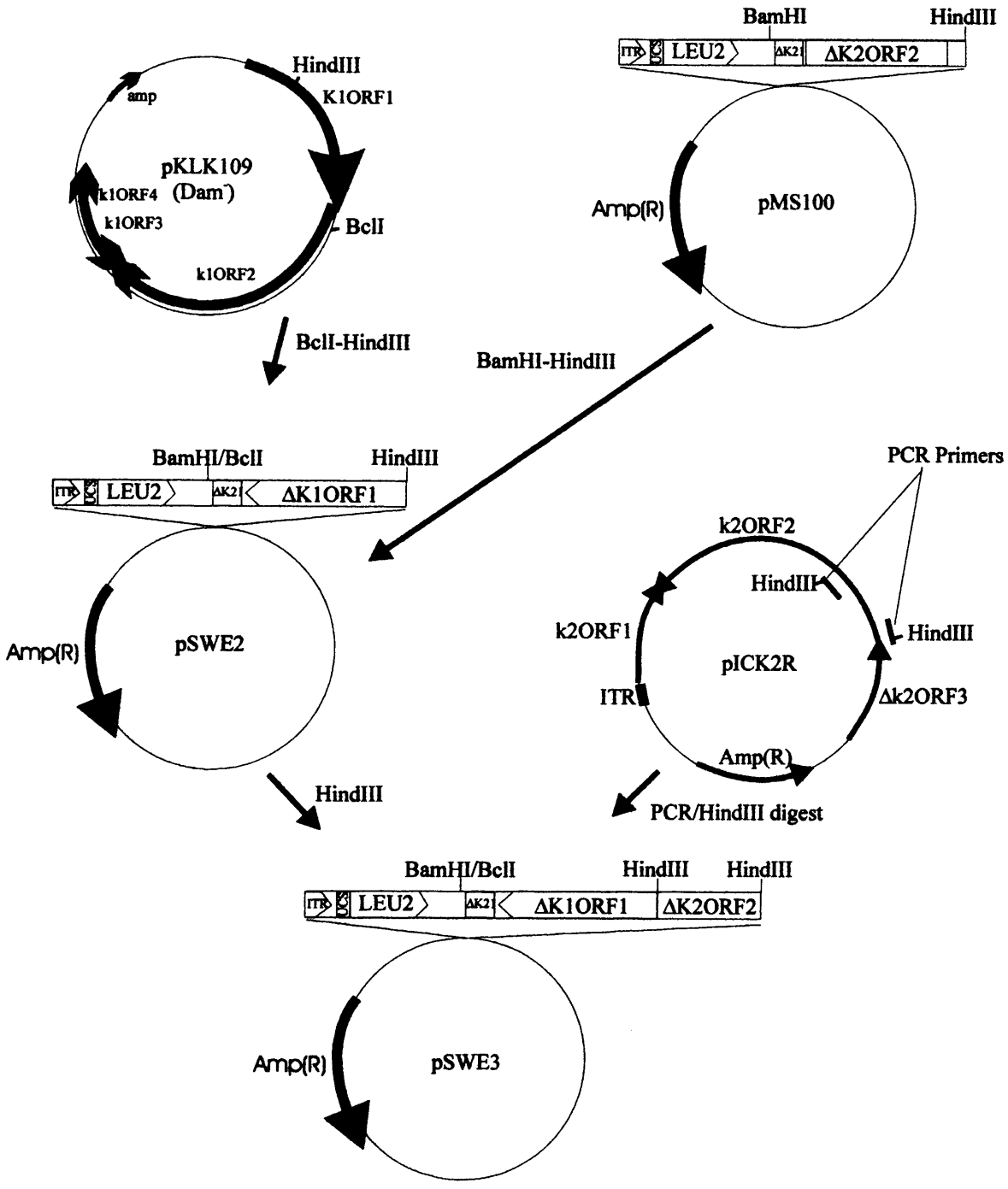
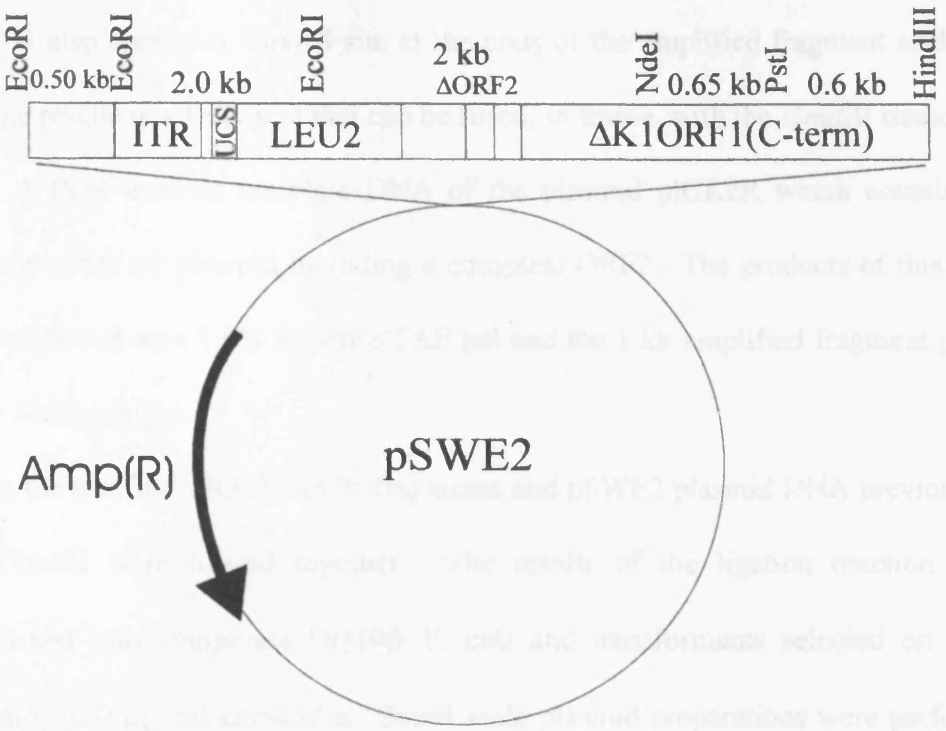


Figure 4.2.1: Construction of the k1ORF1-k2ORF2 hybrid DNA polymerase transformation cartridge.

The next step was to generate a suitable C-terminal fragment of the k1ORF1 gene for further manipulation. Fortuitously, due to the rarity of available restriction sites within the A-T rich linear plasmid DNA, a 2.4kb *Hind*III-*Bcl*I restriction fragment contained the desired central and C-terminal regions of the ORF1 gene.

Digestion of plasmid pMS100 with restriction enzymes *Bam*HI and *Hind*III resulted in a 2.68 kb insert fragment and a 5.32 kb vector fragment. The latter was band-purified from an agarose gel using standard techniques. The 5.32 kb *Bam*HI-*Hind*III pMS100 vector and the 2.4 kb *Hind*III-*Bcl*I pKLK109 fragment were then ligated together using standard ligation techniques with a ratio of insert to vector of 10:1. *E. coli* transformants were selected on Luria agar plates containing 100 μg/ml Ampicillin. Small scale plasmid preparations were performed on a number of transformants until one with the correct recombinant plasmid was found. A large scale preparation of this plasmid DNA was performed and its structure confirmed by restriction enzyme analysis. This analysis showed that, as expected, a single *Hind*III site and a double treatment with *Hind*III and *Pst*I resulted in the production of a 616 bp and a parental band. This confirmed the correct orientation of the Δk1ORF1 fragment within the recombinant plasmid. The recombinant plasmid was designated pSWE2. [see Figure 4.2.2]

The next requirement was to make an in-frame fusion of the N-terminal k2ORF2 gene with the k1ORF1 fragment on the pSWE2 plasmid. Due to the lack of available restriction sites, two primers were designed for the PCR amplification of the k2ORF2 N-terminus. These two primers were designated KFUSH1 and KFUSH2 for the forward and reverse primers, respectively. The 5' primer extends through the extreme N-terminus of k2ORF2 and introduced *Hind*III restriction sequences to facilitate cloning. In addition to their use in the amplification of the 1 kb k2ORF2 N-Terminal fragment, both KFUSH1 and



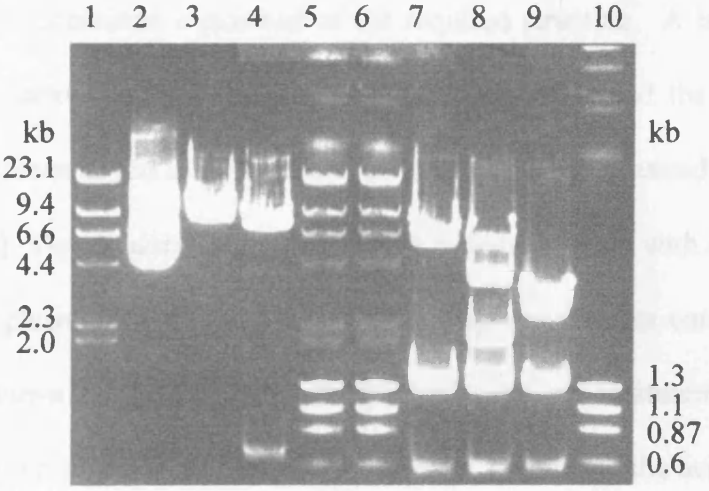
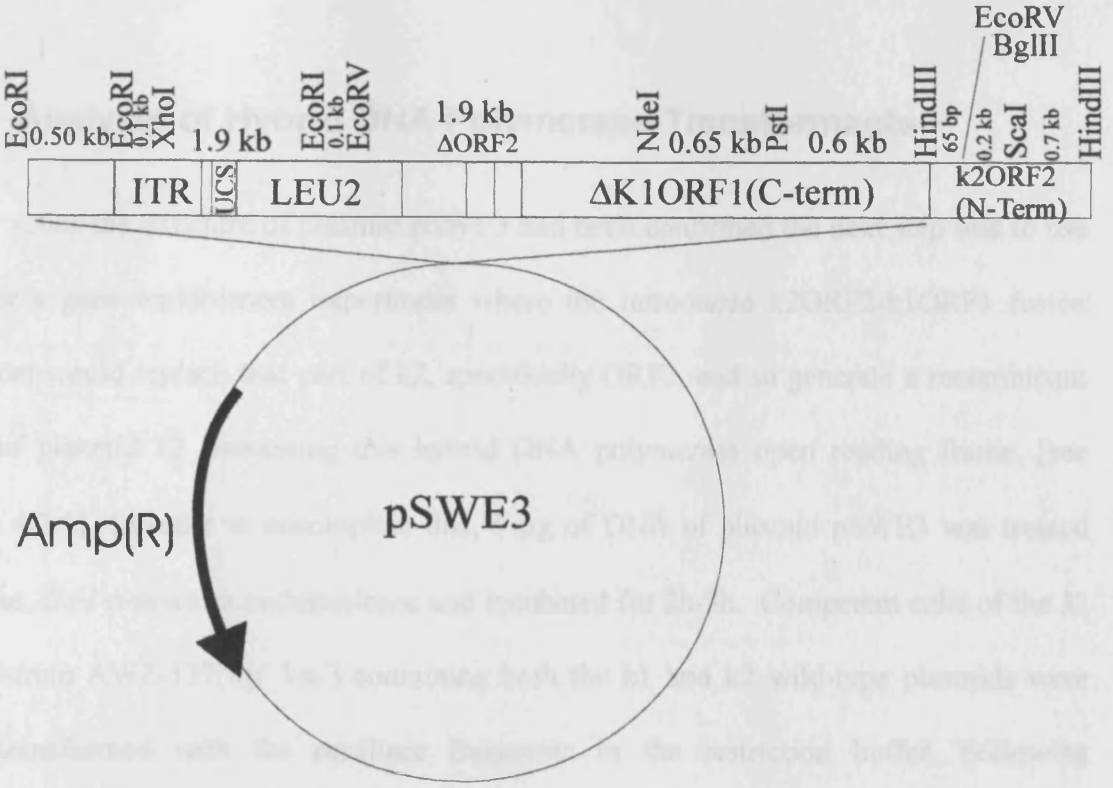


Figure 4.2.2: Plasmid pSWE2 digested with the following restriction endonucleases and samples electrophoresed on 0.7% agarose gel..
Lane 1: λ/*Hind*III molecular weight standards; Lane 2: untreated plasmid;
Lane 3: *Hind*III treated; Lane 4: *Hind*III-*Pst*I treated; Lane 5 and 6: λ/*Hind*III-ϕX174/*Hae*III molecular weight standards; Lane 7: *Eco*RI treated;
Lane 8: *Eco*RI-*Nde*I treated; Lane 9: *Eco*RI-*Pst*I treated;
Lane 10: ϕX174/*Hae*III molecular weight standards

KFUSH2 also append a *Hind*III site at the ends of the amplified fragment such that their cleavage results in a fragment that can be fused, in frame, with the *Hind*III treated k1ORF1 gene. A PCR used as template DNA of the plasmid pICK2R which contains a 5.5kb fragment of the k2 plasmid including a complete ORF2. The products of this PCR were electrophoresed on a 1.0% Agarose/TAE gel and the 1 kb amplified fragment purified for further manipulation.

Finally, the purified k2ORF2 PCR fragments and pSWE2 plasmid DNA previously treated with HindIII were ligated together. The results of the ligation reaction were then transformed into competent DH10β E. coli and transformants selected on Luria agar containing 100 μg /ml ampicillin. Small scale plasmid preparations were performed on a number of colonies resulting from this transformation and restriction analysis of the resulting DNA was used in order to determine if the insert was in the correct orientation. One isolated strain contained a plasmid of the required structure. A large scale alkaline lysis plasmid preparation of this isolate was then performed and the resulting plasmid DNA was restriction mapped in order to ensure that the correct plasmid had been isolated. [see Figure 4.2.3] The sequencing confirmed the expected result, with a HindIII digest of the recombinant plasmid yielding a 1kb k2ORF2 fragment and the correct orientation of the insert was shown by the BglII-PstI double endonuclease treatment giving a 681 bp fragment, in addition to the vector fragment. In order to confirm the accuracy of the PCR fragment, sequencing was also performed on the PCR derived fragment within the plasmid. Following confirmation of plasmid structure, this plasmid was designated pSWE3.



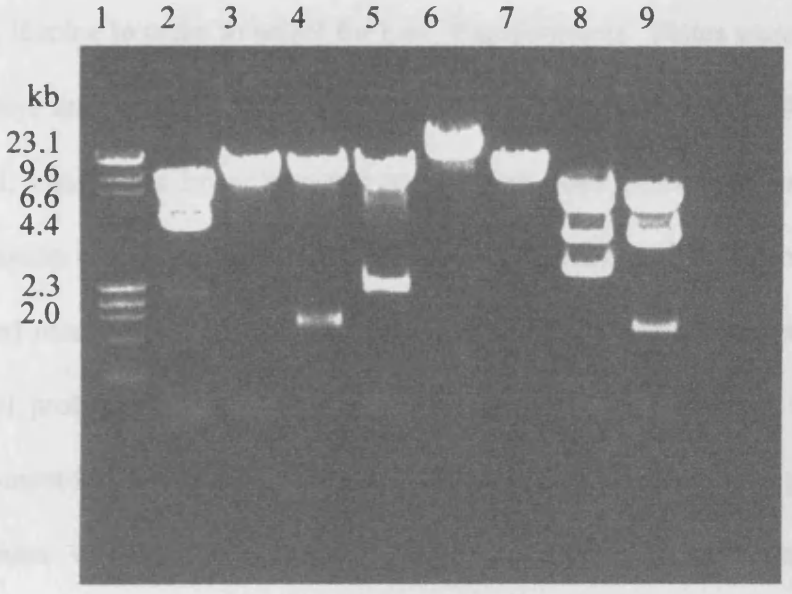


Figure 4.2.3: Restriction analysis of pSWE3, domain swap vector.
Lane 1: λ/*Hind*III DNA molecular weight standards; Lane 2: *Eco*RV treated plasmid; Lane 3: *Pst*I treated plasmid; Lane 4: *Bgl*II - *Pst*I treated plasmid; Lane 5: *Hind*III treated plasmid; Lane 6: *Bgl*II treated plasmid; Lane 7: *Pst*I- *Sca*I treated plasmid; Lane 8: Xho*I*-*Sca*I treated plasmid; Lane 9: *Pst*I - *Eco*RV treated plasmid

4.3 Analysis of Hybrid DNA Polymerase Transformants

Once the structure of plasmid pSWE3 had been confirmed the next step was to use this for a gene replacement experiment where the introduced k2ORF2-k1ORF1 fusion fragment would replace that part of k2, specifically ORF2, and so generate a recombinant form of plasmid k2 containing this hybrid DNA polymerase open reading frame. [see Figure 4.2.1] In order to accomplish this, 5 μg of DNA of plasmid pSWE3 was treated with the *XhoI* restriction endonuclease and incubated for 2h-3h. Competent cells of the *K. lactis* strain AWZ-137(trp^- leu^-) containing both the k1 and k2 wild-type plasmids were then transformed with the resultant fragments in the restriction buffer. Following transformation, the cells were spread onto SD plates supplemented with tryptophan, but lacking leucine in order to select for Leu^+ transformants. Plates were incubated at 30 C for three days and growth data were then analysed. Approximately 500 transformants were isolated; small scale linear plasmid preparations were performed and their linear plasmid composition investigated by both restriction enzyme and Southern blot analysis. The expected result of a Southern blot analysis of LEU+ transformants using a K1ORF1 C-terminal probe would be that two bands would be detected: the wild type k1 and the recombinant k2. Initially no colonies were isolated with this altered plasmid content. This experiment was repeated several times and, finally, two correct transformants were isolated. Their plasmid composition, as expected, consisted of the two wild type plasmids, k1 and k2, and the additional recombinant k2 plasmid. [see Figure 4.3.1]

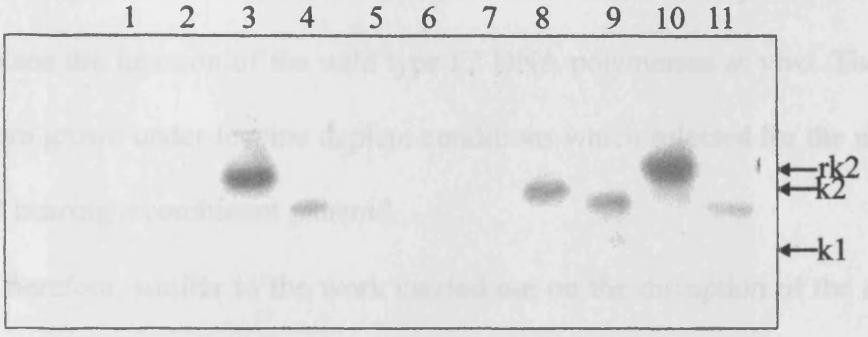


Figure 4.3.1: Autoradiograph of Southern blot analysis of linear plasmid species from LEU+ transformants of strain AWZ-137 after transformation with pSWE3 transformation cartridge. Total DNA samples were electrophoresed on a 0.7% agarose gel supplemented with 0.1% SDS, blotted onto a nylon membrane and hybridized to *Saccharomyces cerevisiae* LEU2 gene probe.
Lane1: AWZ-137 parental strain.
Lane 2 to 12: individual transformants from the Southern blot screen.
Recombinant k2 plasmids can be seen in lanes 3 and 10.
In lanes 4, 8, 9 and 11 a second type of LEU2 containing species is present which seems to correspond to neither the k2 or k1 plasmids, thus suggesting it may be some type of circular linear plasmid species which has formed and may replicate via the fortuitous ARS sequences contained within the linear plasmids.

4.4 Analysis of Hybrid Transformants

Following the isolation of the two transformants containing the recombinant k2 plasmid, an experiment was performed to ascertain whether the hybrid DNA polymerases could replace the function of the wild type k2 DNA polymerase *in vivo*. The recombinant clones were grown under leucine deplete conditions which selected for the maintenance of the LEU2 bearing recombinant plasmid.

Therefore, similar to the work carried out on the disruption of the Δk1ORF1 (see Chapter 3) it was reasoned that growth under this selective regimen would lead to the eventual displacement of the wild-type k2 by the recombinant k2 if the k2ORF2-k1ORF1 recombinant DNA polymerase was able to substitute for the k2ORF2 product (see Chapter 3). Repeated sub-culturing experiments were performed in order to determine the functionality of the hybrid DNA polymerase. Following 20 sub-culturing stages, the equivalent of 250 generations, detailed analysis was performed in order to determine the plasmid composition of the culture.

Thus the recombinant clones were sub-cultured in SD medium supplemented with tryptophan for three days for each of the twenty rounds of sub-culturing and linear plasmid DNA purified and subjected to agarose gel electrophoresis. Southern blot analysis was then carried out on the resultant plasmid DNA using a *Saccharomyces cerevisiae* LEU2 probe. These experiments showed that the wild type plasmid was still present along with the recombinant form. Moreover, the recombinant k2 species represented only a small proportion of the plasmid content compared to the wild type k2, as judged by visual inspection of Ethidium Bromide stained agarose gels. Figure 4.4.1 shows the results of restriction analysis carried out on plasmid DNA purified using the Ammonium Sulphate purified linear plasmid DNA from isolates containing the recombinant k2 plasmid.

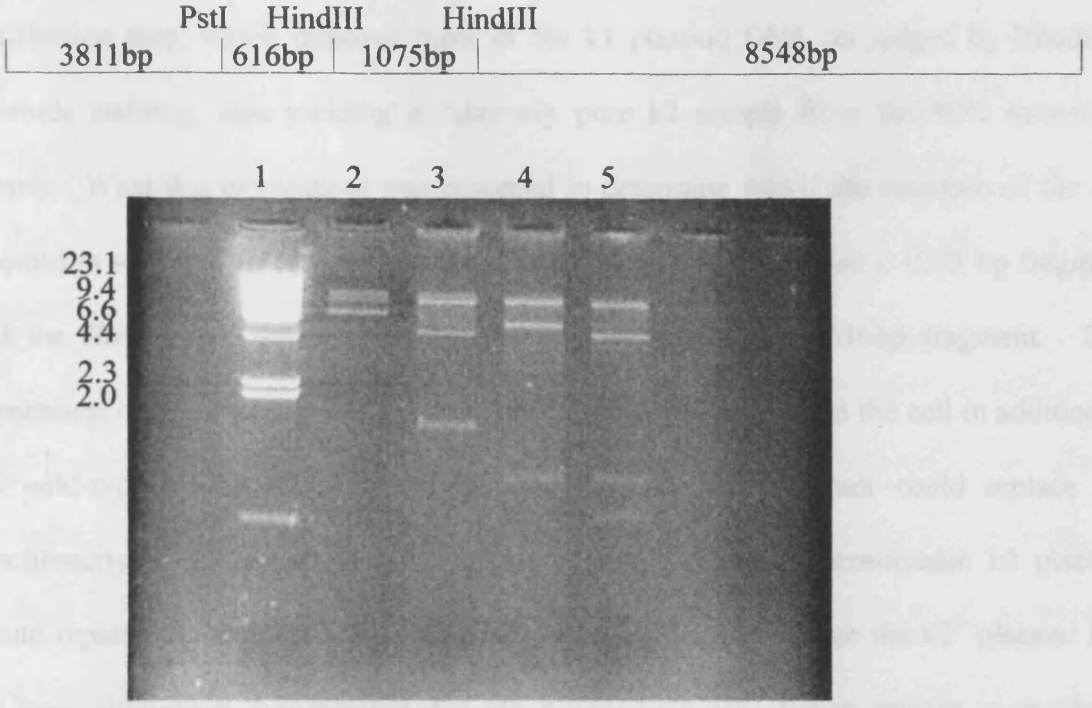


Figure 4.4.1: Restriction map of the linear plasmid content of a leucine prototrophic pSWE3 catridge transformant. These digests were used to ensure that the correct k2ORF2-k1ORF1 hybrid plasmid structure had been obtained.
Samples treated were the linear plasmid fraction of a leucine protrophic isolate purified using the ammonium sulphate method described in this study, performed so as to minimize k1 contamination.
Lane 1: l/*Hind*III-fX174/*Hae*III molecular weight standard. Lane 2: Untreated plasmid DNA. Lane 3: *Pst*I treated plasmid DNA. Lane 4: *Hind*III treated plasmid DNA.
Lane 5: *Hind*III-*Pst*I treated plasmid DNA.
The results are as expected (see restriction map), thus demonstrating that a correct rk2 construct was contained within the transformant. This proved that the rk2 plasmid was unable to replace the native k2 plasmid *in vivo*.

The Ammonium Sulphate preparation was performed with an 80% and 85% saturation purification step, which removes most of the k1 plasmid DNA, as judged by Ethidium Bromide staining, thus yielding a relatively pure k2 sample from the 90% saturation sample. What this experiment was designed to determine was if the structure of the rk2 plasmid *in vivo* was as expected. The *Hind*III digest clearly yielded a 1075 bp fragment and the *Hind*III-*Pst*I double endonuclease treatment yielded a 616bp fragment. This experiment demonstrated that the correct rk2 plasmid was present in the cell in addition to the wild-type k2 plasmid. What was expected, if the construct could replace the functionality of the native k2 DNA polymerase, was that the recombinant k2 plasmid would replace the native k2 plasmid under leucine selection. Since the $k2^+$ plasmid had not been eliminated, I concluded that the k2ORF2-k1ORF1 fusion protein is unable to substitute for the native k2ORF2 protein and therefore the wild-type k2 plasmid was required for maintenance of the recombinant species. It did, however, demonstrate that the recombinant plasmid was being maintained within the cell and that its structure was as expected.

4.5 Discussion

The successful, in frame fusion between the k2ORF2 and k1ORF1 genes, followed by the introduction of the fusion protein into the k2 plasmid demonstrated that the terminal recognition domain does not appear to be the sole point of the DNA polymerase protein responsible for plasmid specificity within the k2 system.

The inability of the k2ORF2-k1ORF1 hybrid DNA polymerase to replace the function of the normal DNA polymerase must suggest one of several possibilities. Firstly, the construct contained insufficient N-terminal amino acids to constitute a fully functional domain; secondly, the domain structure of the DNA polymerase may be more complex than proposed, with multi-domain, tertiary and quaternary structural interactions playing an essential role in defining plasmid specificity e.g. terminal recognition; thirdly, the central and/or C-terminal region could have an important role in terminal recognition. In order to address these possibilities, more research is necessary, but due to current limitations in techniques used for this type of manipulation, an exhaustive experiment involving a large number of possible sized fragments would be very time consuming, although it is possible. The possibility that the C-terminal region is important in terminal recognition is unlikely, based on a comparison with the adenovirus system, but a sequence specific interaction between the two cannot be ruled out.

Chapter 5: Discussion

5.1 Introduction

As a conclusion to this work I would like to recap the main points of this study and draw some conclusion, as well as make suggestions for further experiments which may help to clarify the still murky picture of the *K. lactis* "killer" plasmid system.

5.2 Biochemical Analysis of the k2ORF2 Putative DNA Polymerase

This part of the study set out to examine the possibility that the k2 terminal protein, earlier detected biochemically, was encoded by the cryptic amino terminus of the putative k2ORF2 DNA polymerase. The immunochemical approach used to address this question was chosen because it was expected to be sensitive, reproducible and reliable. Once an antibody is available for a particular protein it allows sophisticated protein and gene expression studies to be carried out as well as giving a means of very specific purification.

The inability of the antibody raised against the GST-k2ORF2 to detect k2 linear plasmid terminal proteins could stem from several root causes. The two most likely possibilities are firstly that too little terminal protein was present so that the signal was below the level of detection, or, secondly, that many of the terminal proteins attached to the plasmids, were in a degraded state caused by release of proteases during the purification procedure. We believe that the terminal proteins do remain largely intact following isolation, because of the retarded migration of linear DNA fragments detected during gel electrophoresis and due to I^{125} detection studies. (Kikuchi *et al*, 1984; Stam *et al*,1986; Reay, 1997)

The amino terminal sequence evidence provided by Takeda *et al*, while implicating k2ORF2 as encoding the terminal protein is ambiguous about where the terminal protein actually starts. (Takeda *et al*, 1996) Figure 2.5.1 illustrates the terminal protein and DNA polymerase domains which would be predicted using the sequence data obtained. These sequence data were likely to be derived from within the terminal protein since calculation of the size of the terminal protein predicts that it would span some conserved DNA polymerase domains, an unlikely situation.

The third possibility is that the terminal protein is folded in such a way as to sterically hinder the binding of antibody to the extreme N-terminal domain used for antibody production as only partially-denaturing conditions were used to analyse the protein-plasmid complexes.

One experiment which might overcome many of the difficulties encountered would be to repeat the raising of antibodies using a peptide 3-4 times the size. This synthetic route is costly but since previous attempts to express large fragments of k2ORF2 have been unsuccessful, it appears the only available method. (Ambrose, 1993) The peptide used in this experiment was constructed by producing a oligonucleotide encoding the amino terminus of the k2ORF2 protein using *E. coli* codon bias. Constructing a larger fragment would require the synthesis of several oligonucleotides, which could then be ligated together to form the larger fragment. The use of this larger fragment might also solve much of the sensitivity problems encountered, as a much higher titre of ORF2 specific antibodies would be present within the antibody sample and would limit the effect of steric hindrance as more of the protein would be available for interaction with the antibodies. This experiment could be practically done but the cost may be prohibitive! A second possibility is that antibodies could be raised to a number of small epitopes along the amino terminus of the k2ORF2 protein.

5.3 Genetic Analysis of the k1ORF1 DNA Polymerase

This series of experiments was designed to examine whether the k1 putative DNA polymerase encoded by k1ORF1 was essential for k1 maintenance or if its function could be carried out by the k2ORF2 encoded DNA polymerase. It was important to establish whether the k1 and k2 plasmids can utilise the same DNA polymerase for replication.

In evolutionary terms, the sequence and functional diversity of the k1 and k2 putative DNA polymerases suggests a co-operative relationship between k1 and k2. The k1 plasmid encodes a DNA polymerase but the rest of its compact genome is entirely given over to killer toxin production and resistance. Plasmid k2, which also has a DNA polymerase, appears to encode no components of the "killer" toxin system, but must encode the transcriptional and replicative functions needed by the plasmids. The dependency of k1 upon k2 is clear but what advantage the relationship with k1 confers on k2 is still unclear. In comparison the *S. kluyveri* plasmid, pSKL, which is very similar to k2, exists alone. (Hishinuma and Hirai, 1991)

There are at least two possible types of interaction which could exist between k2 and k1. The first is that k1 confers a competitive advantage to the *K. lactis* cell in relation to non-killer yeast. Thus natural selection has helped the development of a stable k1-k2 interaction in which selection drives the formation of a stable binary plasmid system. The second is that k1 is efficient enough at replicating that it can maintain itself in the cell in the absence of any positive selection, but not so efficient as to inhibit the replication of k2 by using too many of the k2 replication proteins for its own duplication. I believe that there is also a third alternative, which is another shade of the first possibility: at some point k1 must have either entered a cell through a mating or fusion or arisen through a series of alterations to the k2 plasmid and insertion of killer toxin DNA. Following its introduction

k1 produced the killer toxin and resistance proteins and hence provided a form of negative selection for its loss. Any cell spontaneously losing the k1 plasmid was affected by the toxin. This, over time, may have led to the development of a stable system. If linear plasmid experimental technology was better developed, it would be interesting to assay the relative quantities of replication or transcription proteins produced by pSKL and k2, as the relationship with k1 that exists in the *K. lactis* system may have selected for elevated levels of these proteins. The suggestion is therefore that the pSKL plasmid may be able to carry out a similar function as k2 by supporting k1 replication. This experiment could be performed by performing a protoplast fusion between *K. lactis* cells containing k2 and k1 and *S. kluyveri* cells containing pSKL. The k2 plasmid would need to bear some selectable marker which could be used to counter-select for its maintenance thus showing whether pSKL can carry out its k1 maintenance functions.

The experiments performed demonstrated that k1ORF1 is essential to k1 maintenance and that the k2 DNA polymerase is unable to complement a defect in this gene, thus confirming that both DNA polymerases are required. Why two DNA polymerases are required is perhaps best explained by the idea the k1 plasmid was introduced into the cell after k2 had already become established. This means that both plasmids had their own DNA polymerase and because their starting points were dissimilar and hence could not easily converge. Also, the terminal protein domains on the respective DNA polymerases must have evolved to optimally recognise their inverted terminal repeat regions. The co-evolution of ITR and terminal protein must be a pre-requisite for this type of system to function. Historically the proteins which would become the terminal proteins must have been some type of generic charged protein which was capable of binding DNA and became capable of forming the covalent attachment vital to protein primed replication. The ability of the terminus to adapt through random mutation and maintain the high

binding affinity of the terminal protein would have been vital not only to replicative efficiency but also for replication itself. During the continuing co-evolution, any change in the make-up of the ITR which stimulated binding resulted in an evolutionary advantage whilst any which inhibited binding has a disadvantage.

Following on from the previous point, a second aspect of the terminal protein-plasmid relationship which remains unanswered is whether the terminal protein-DNA polymerase hybrids are "single use". The DNA polymerase and terminal proteins appear to be produced as a single polypeptide and during initiation of DNA synthesis the terminal protein forms a covalent attachment to the terminus of a linear plasmid. The DNA polymerase, however, needs to be able to process its way through the DNA strand in order to function. There are two ways this can happen: either the terminal protein and DNA polymerase become separated during initiation, leaving the DNA polymerase free to process the DNA strand, but therefore resulting in the destruction of the DNA polymerases ability to initiate replication, or the DNA polymerase draws the DNA strand through itself. The first possibility seems the most likely, both because the second option would be thermodynamically unfavourable and secondly because the size of the terminal proteins is known. How the terminal protein separates from the DNA polymerase is unknown, but could occur in three different ways: the first is that a generic yeast protease is responsible for this, the second is that the linear plasmids encode their own specific protease to carry out this function and the third is that the terminal protein-DNA polymerase protein can carry out this function itself. The first option is cast into some doubt because of the relatively wide host range of the k1-k2 linear plasmids. No real evidence is available either for or against the other two options. If the single use enzyme theory is true, then this may be the main means through which the linear plasmids regulate their copy number. The DNA polymerase - terminal protein complex is known to be transcribed at a very low

level, so it is conceivable that this is one of the main factors affecting copy level. In order for this to function efficiently a positive feedback loop can not be established. Following any transient increase in level of this enzyme a rapid increase in replication could be envisaged, although the kinetics of such a effect would need to be such that the enzymes were rapidly used up. The mechanism by which it is used up and does not cause any long term effects on the linear plasmid copy levels could be through its "single use" nature, whereby each DNA polymerase-Terminal protein polypeptide can only drive one replication event.

Further DNA analysis that could be done using the gene replacement technique would be to attempt to replace the various function of domains of the k2 DNA polymerase with those of the pSKL DNA polymerase in order to examine how closely related these two plasmids are. This would be a "domain swap" experiment as carried out for the k1 an k2 plasmids. If this experiment was successful it could also allow the prediction of which amino acids and structural elements are required for binding specificity in the terminal protein.

5.4 Production and Characterisation of a Hybrid DNA Polymerase

This phase of the study attempted to localise the site of the plasmid recognition domain to either the putative terminal protein or DNA polymerase regions. The theory was that because the polymerase and exonuclease domains of the k1ORF1 and k2ORF2 DNA polymerases are so similar it is likely that they may be generic polymerases which could replicate either k1 or k2 plasmids. However the divergent N-terminal regions would target each of them to their specific plasmid. Evidence that the k2 terminal protein is encoded by the amino-terminal domain of the DNA polymerase also strengthens the likelihood of some targeting feature of the amino-terminal domain as the terminal protein

is normally central in the initiation of DNA synthesis in protein primed genomes. A hybrid DNA polymerase was therefore constructed with a k2 amino terminal protein fragment and k1 central and carboxyl terminal DNA polymerase domains. If this construct could replace the normal k2 DNA polymerase it would mean that the k2 terminal specificity determinant had been successfully linked to the k1 DNA polymerase domains hence constructing a pseudo k2 DNA polymerase. Moreover it would imply that the DNA polymerase parts of the 2 proteins were functionally redundant.

The apparent inability of the hybrid k1-k2 DNA polymerase to substitute for the wild type k2 DNA polymerase and terminal protein pair is a complex issue and could have stemmed from a variety of causes. It is possible that despite sharing a common functionality the DNA polymerase domains have evolved to interact optimally with their attached terminal proteins. This could render the hybrid DNA polymerase reaction much less favourable, thus giving a replicative advantage to plasmid using the native wild type k2-DNA polymerases. Thus the hybrid DNA polymerase might be functional but at a less efficient rate making it unable to compete with the more efficient wild type k2-DNA polymerase. Imposition of the $Leu2^+$ selection would select for the maintenance of the hybrid DNA polymerase- containing plasmid, but if the hybrid polymerase had a much lower replication rate then plasmid carrying that gene may have difficulty becoming dominant. The extreme situation is that the hybrid DNA polymerase is either not expressed properly, or does not function at all in DNA replication. Both of these possibilities fit the evidence and the continuing presence of the hybrid DNA polymerase containing plasmid could be the result of replication brought about by wild type k2 DNA polymerase. For this experiment, a temperature sensitive mutation in the k2 DNA polymerase would be ideal.

In order to determine if the hybrid DNA polymerase is function we would need a way of removing the wild type k2 DNA polymerase from the cell. One possible way to do this would be to construct a double mutant. One mutant would contain the wild type k2ORF2 DNA polymerase as well as bearing a counter-selectable marker on a non-essential gene, i.e. k2ORF1. (Schaffrath *et al*, 1992) The second would contain the hybrid DNA polymerase bearing k2 plasmid as constructed here, bearing the LEU2 gene. Counter selection for a marker contained within the k1ORF1 disruption version of k2 could be used to drive the loss of the plasmid from the cell. If the hybrid DNA polymerase is functional isolates will be found that lack wild type k2. If the hybrid DNA polymerase is not functional no pure hybrid k2 isolates will be found.

This study has also described a simplified method which utilises a standard ammonium sulphate precipitation procedure to facilitate rapid purification of linear plasmids (see Materials and Methods). Further refinement of this procedure could lead to an easy method of purification which would allow the purification of large quantities of linear plasmid material for experimentation. The improvements which could be made to it would include definition of the optimum pH and salt concentration for precipitation. Currently, preliminary evidence obtained during this study has shown that different salt concentrations precipitate the k1 and k2 plasmids differentially, as would be expected when one considers that ammonium sulphate precipitation functions through the neutralisation of surface charges, thus allowing aggregation. This method was based on that described by Reay (described in Materials and Methods) in terms of the conditions the yeast cells are grown in and the means of cellular disruption, but whereas Reay made use of charge and size separation chromatography, this method makes use of Ammonium Sulphate in order to selectively precipitate the linear plasmids. (Reay, 1997) The main advantage of this method over those earlier described is the speed at which isolation can be

done. (Reay, 1996; Stam *et al*, 1986) Whereas the other methods described require at least two days, a similar isolation can be done in about 4 hours. The quality of material seems to be similar to that of Reay although detailed comparison has not been carried out. The terminal proteins also appear to be intact, judging by the behaviour of the linear plasmids in gel electrophoresis. The amount of material obtained appears to be approximately 10-20% more that that obtained by Reay from the same number of cells.

What is really needed in order to better analyse the system is a new method of genetic analysis which greatly simplifies the manipulation of linear plasmid DNA and the development of a good biochemical expression system to produce the individual plasmid open reading frame products in a pure form. One candidate system is Baculovirus, which has been used previously in this laboratory to express a recombinant k2ORF5 form. (Schaffrath and Meacock, 1995) Baculovirus needs to be more fully investigated as a linear plasmid ORF expression system, although it is currently believed that it is limited to the expression of small ORF products. If these proteins were available in pure forms detailed work could be carried out on them similar to that already performed on the Adenoviral and Phi-29 systems studying the linear plasmid replication *in vitro.* (Caldentey *et al*, 1993; Pronk *et al*, 1992) Out of this would come detailed knowledge about the initialisation of linear plasmid replication as well as detailed knowledge about the function of the k2 RNA polymerase and allow characterisation of the other k2 open reading frames for which no current function is known.

5.5 General Discussion

The preceding sections summarised the different sections of this study and possible alternative and/or supplementary experiments. In conclusion I feel it is useful to review again our current state of knowledge about the replication of the k1 and k2 plasmids.

It is believed that the linear plasmid replication initiates from plasmid specific terminal proteins and proceeds uni-directionally, without Okazaki intermediates. Also, it is known that the terminal protein and DNA polymerase are transcribed and presumably translated mono-cistronically, at a low level. This means that k1ORF1p and k2ORF2p are most likely expressed and bind the corresponding plasmid terminus during initiation the terminal protein and DNA polymerase domains become separated by proteolytic cleavage. The presence or nature of accessory factors is currently unknown, although helicases would presumably be needed in order to unwind the DNA during replication, as well as a protease to perform the TP-DNA polymerase separation. K2ORF4 is believed to be a good candidate for the helicase role based on its similarity to the vaccinia virus DNA dependent ATP helicase. (Wilson and Meacock, 1988; Soond, 1994) The similar vacinnia product is believed to unwind the vacinnia linear DNA duplex in order to allow transcription to occur.

The family of linear plasmids has grown greatly since the original discovery of the *K. lactis* killer plasmids and several, including the killer plasmids have been relatively well characterised. Many have had their complete sequences resolved. Analysis of these sequences has, apart from demonstrating that most linear plasmids encode their own transcription/translation system, left key questions unanswered about why this class is so widespread. At first these plasmids were thought to be something of an oddity, but it is becoming apparent that this family of genetic entities is highly successful and has spread

throughout the plant and yeast phyla, although with, so far, no good examples of higher animals hosting them.

Why these entities are so successful is a mystery, especially in terms of nice, clear, Darwinian evolutionary thinking. Apart from a few examples, including the killer plasmids and the *kalilo* senescence plasmids of *Neurospora intermedia*, most of these plasmids have been assigned no clear phenotype; they appear to give no conceivable advantage to the host. These plasmids could be best described as genetic parasites, or "selfish DNA", living off a rich genetic resource, but unlike many parasites, the linear plasmids appear to be, as a class, non-harmful due to the tight control of their replication processes most likely imposed by the inefficiency of their promoter RNA polymerase recognition system. This is perhaps the key to their success. Perhaps they are a prime example of a type of neutral selection in which a non-harmful parasite which confers no selective advantage but is simply efficient and good at reproducing itself has survived and prospered. The linear plasmids must, however place some degree of strain on the host system, this implies that yeast and plant cells have evolved so that they have a metabolic over-capacity which can be exploited by parasites.

The origins of the linear plasmids is also shrouded in mystery. Due to their apparent viral nature many have theorised that they stem from viruses, but none have been found to have clear remnants of viral capsid protein genes. Unless all of the linear plasmids stemmed from a single event long ago, one could expect clear examples of viral gene expression to be present, other than that required for simple replication. Structurally, the *Morchella conica* pCM32 plasmid has been reported to share some structural similarities with simian adenovirus but whether this means that it originated from adenovirus or it has acquired some of the DNA by a recombination event is still unclear. (Rohe *et al*, 1991) It could have also evolved from a common ancestor. If all linear

plasmids in fact stem from viruses, it perhaps is not so much of a mystery as to why they have lost the structural genes over time if one presumes a constant pressure for compactness so allowing rapid replication times and hence more effective competing for resources. The fact, however that all linear plasmids appear to be in a similar evolutionary state is interesting because this perhaps mean that some event in history triggered this to occur, or it could mean that linear plasmids and viruses diverged from each other, taking opposite tacks in how to approach replication. Viruses perhaps acquiring protective coatings to allow them to move more easily from cells to cell, but the linear plasmids remaining stable in a line once established. Either way, I am not convinced that the linear plasmids arose simply through a random event which caused the viruses to lose their normal form of replication. Perhaps this type of replication is preferable in relatively isolated environments in which other members of the same species are rare and growth rates slow, but many viruses have adapted to similar conditions and been successful. The resolution of this mystery will require so very detailed comparison of viral and linear plasmid sequences. This analysis should shed much light on the inter-relationship between linear plasmids and viruses.

Chapter 6: Materials and Methods

6.1 Strains and Media

6.1.1 E. coli Strains

NM522: [Δ(lac proAB) thi supE Δ(hsdMS-mcrB) 5 r_K-m_K- mcrB/ F' proAB+ $lacI^Q$ lacZ Δm15]; Gough and Murray (1983)

6.1.2 Kluyveromyces lactis Strains

AWZ-137: *leu2, trp1* [$k1^+$ $k2^+$] Kämper *et al.* (1991)

NK-40: α, ade1, ade2, leu2, [$k1^0$ $k2^+$]

IFO1267: Wild Type [$k1^+$ $k2^+$] Gunge *et al.* (1981)

SD801: a, trp1, lac4 [$k1^0$ $k2^0$] Stark & Milner (1989)

SD802: a, trp1, lac4, [$k1^0$ $k2^+$] Wilson (1988)

6.1.3 Growth Media

Luria Broth: All growth of *E. coli* was done in Luria Broth (10g bacto-tryptone, 5g bacto-yeast extract, 10g NaCl and pH adjusted to 7.5 with NaOH or HCl per litre) Luria Broth was autoclaved and stored in sealed bottles or flasks at room temperature until used.

Synthetic Defined Media: Made in 500ml batches by dissolving 3.5g Yeast Nitrogen Base lacking amino acids in 500ml of distilled water and. The resulting solution was autoclaved and then stored in sealed bottles until use. Prior to use medium was supplemented with glucose to a final concentration of 2% and with amino acids as appropriate. All disruption

analysis of linear plasmids in *K. lactis* was carried out using this media, supplemented with the following amino acids as required at 30C: L-Leucine (30 mg/L), L-Tryptophan (20 mg/L). Solid media was supplemented with 1.5% (w/v) Difco purified agar.

Yeast Peptone Broth: 2% peptone, 1% yeast extract. Made in 400ml volumes, autoclaved and stored in sealed bottles or flasks at room temperature until used. Before use YPD was supplemented with glucose to a final concentration of 2%. Rapid growth of large quantities of *K. lactis* for linear plasmid purification was carried by growing the required strains in this media at 30C.

6.2 Stock Solutions

Ampicillin Stock Solution: A 100 mg/ml stock solution was made in water and filter sterilized. 400 µl aliquots were then stored at -20C. This stock solution was diluted to a final concentration of 50-100 µg/ml in growth media.

6.2.1 Stock Solutions for DNA Analysis

Church-Gilbert Hybridisation Buffer: 0.5M NaH_2PO_4/Na_2HPO_4 (pH 7.4), 1mM EDTA, 7% SDS (w/v) (Church and Gilbert, 1984).

DNA Loading Buffer: The 6x loading buffer used consisted of 30% glycerol and 0.25% bromophenol blue

DNase Free RNase: A 10mg/ml solution was made by dissolving lyophilised RNaseA in distilled sterile water. The solution is then rapidly heated to 100C for 20 minutes and allowed to cool slowly to room temperature. The solution was then divided into 10µl aliquots and store frozen at -20 C.

ϕX174\HaeIII standard: This marker was diluted to a concentration of 1µg\5µl, 5µl was routinely used per gel. DNA fragment sizes were: 1,353bp, 1,078 bp, 872 bp, 603 bp, 310 bp, 281 bp, 271 bp, 234 bp, 194 bp, 118 bp, and 72 bp.

λ\Hind III standard: The λ standard used was a 0.5µg\ 5 µl stock made by diluting a 0.5µg\µl stock. Aliquots of 5µl were normally used per gel. DNA fragment sizes were: 23,130 bp, 9,416 bp, 6,557 bp, 4,361 bp, 2,322 bp, 2,027 bp, 564 bp and 125bp.

Alkaline Lysis Plasmid Preparation Solution I: 25mM Tris(8.0), 10mM EDTA (pH 8.0), made up to 100ml with distilled water and autoclaved.

Alkaline Lysis Plasmid Preparation Solution II: 0.4M NaOH, 1% SDS, prepared fresh before use.

Alkaline Lysis Plasmid Preparation Solution III : 60ml 5M potassium acetate, 11.5ml glacial acetic acid, made up to 50 ml with distilled water and autoclaved.

SSC (20x): 0.3M Sodium Citrate (pH 7.0), 3M NaCl

Tris-Acetate EDTA (TAE) Electrophoresis buffer[10x stock]:48.4g Trisma base, 11.42ml glacial acetic acid, 20ml of 0.5M EDTA (pH 8) {Per Litre}

Tris-Borate EDTA (TBE) Electrophoresis Buffer[10x stock]: 108g Trisma base, 55g Boric Acid, 0.5M EDTA(pH 8) 40ml {Per Litre}

6.2.2 Stock Solutions for Protein Analysis

Coomassie Blue Protein Stain/Destain: 50% methanol, 10% acetic acid in distilled water. For the stain 0.375 g Coomassie brilliant blue was added before the addition of the distilled water to 1L.

Lysozyme Solution: 10 mg/ml stock was produced by dissolving lyophilised lysozyme in distilled water and filter sterilised. The solution was then aliquoted into 50 µl amounts and frozen at -20C.

PAGE Cracking Buffer (2x): 60 mM Tris-HCl (pH 6.8), 1% β-mercaptoethanol, 1% SDS, 100% glycerol to give a final concentration of 10%, 0.01% bromophenol blue, made in distilled water.

PAGE Sample Buffer (2x): Twenty milliliter batches of sample buffer were made by combining 0.92g SDS, 2ml β-mercaptoethanol, 4.0g glycerol, 0.3g Tris base, 2ml (0.1[w\v] solution) Bromophenol blue dissolved in 20 ml dH_2O and adjusted to a pH of 6.8 with HCl.

Protein Molecular Weight Standards: Protein molecular weight standards 2,850 Da - 43,000 range (Gibco). Batches of protein standards were prepared for use according to the

manufacturers instructions. 5 µl -10 µl were used per well in mini-protein gels and 20 µl - 30 µl were used for large acrylamide gels. Standards solutions were boiled for 3 minutes before use.

Resuspension Buffer (Inclusion Body Prep): 10mM NaCl, 1mM EDTA (8.0), 50 mM Tris (8.0)

Running Buffer (RB)[10x stock]: 0.5M MOPS, 10mM EDTA, 10mM EGTA, dissolved in distilled water and adjusted to a pH of 7.5 with Acetic acid. Solution was then filter sterilised before use. [This stock is the base for the 0.3M NaCl, 0.4M NaCl and 1.0M NaCl buffers used in killer plasmid purification]

SED Buffer: 1.2M Sorbitol, 20mM EDTA(pH 8.0), 50mM Dithiothreiotol (DTT) dissolved in distilled water and filter sterilised before use. DTT was added immediately before use.

SC Buffer: 1.2M Sorbitol, 0.4M $CaCl_2$ dissolved in distilled water and filter sterilised before use.

SDS-PAGE Electroblotting Buffer[4L]: 12.11 g Trisma Base, 57.65g Glycine, 1L methanol, made up to 4L with distilled water.

SDS-PAGE Electrophoresis buffer: 0.025 M Tris.HCl, 0.192 M Glycine, 0.1% SDS

Thrombin Digestion Buffer: 2.5 mM $CaCl_2$ in Wash Buffer (50 mM Tris (pH 7.5), 150 mM NaCl)

Tris Buffered Saline (TBS)[1L of 10x Stock]: 90g NaCl, 60g Trisma base, pH to 7.9 with HCl.

6.2.3 Stock Solutions for Linear Plasmid Analysis

Linear Plasmid Loading Buffer: Standard DNA loading buffer supplemented with 1% SDS

Micrococcal Nuclease Buffer (10x): 500 mM NaBorate (pH 8.8), 50mM NaCl, 25 mM $CaCl_2$.

Micrococcal Nuclease: 1 mg/ml micrococcal nuclease in 50% glycerol. Aliquoted to 20 µl aliquots.

6.3 DNA Techniques

6.3.1 General Techniques

Agarose gels: The agarose used was Sea-Kem standard agarose. All gels were 0.7% agarose made in 1x Tris Acetate EDTA (TAE) buffer and electrophoresed in the same buffer

Automated DNA Sequencing Automated cycle sequencing was carried out according to the standard ABI automated sequencing method, using a university ABI automated sequencing machine.

DNA Dideoxy sequencing Sequencing was carried out using the Pharmacia T7 sequencing kit according to the manufacturers instructions.

Ligations All ligations were carried out using T4 ligase according to standard methods

Polymerase Chain Reaction (in 20 µl volume): The following were mixed together: 1 µl dNTPs (25 mM), 1 µl $MgSO_4$, 2 µl 10x Vent Buffer, 0.5 µl VENT polymerase, ~ 1 µl (approximately 100-500 ng/ 1 µl) template, 2 + 2 µl of primers, made up to 20 µl with distilled water. All PCR reactions were carried out according to manufacturers recommendations and based on methods described by Ehrlich (1989).

Primer Purification: Purification of all oligonucleotide primers used for both sequencing and PCR was carried out according to the following method: 100 µl of primer was transferred to a fresh microcentrifuge tube and centrifuged for 5 minutes at 10,000 rpm in a to remove debris. The supernatant was then transferred to a fresh microcentrifuge tube. 1 ml n-butanol was then added to the supernatant and mixed. The sample was then centrifuged for 5 minutes at 10,000 rpm. The supernatant was removed, air dried and the pellet resuspended in 10 µl of TE.

Random Prime labeling of DNA: DNA probe labelling was according to the random hexanucleotide method described by Feinberg and Vogelstein (1983). Approximately 50 ng restriction enzyme treated and purified DNA probe, was made up to a volume of 10 µl and boiled for 5 minutes. The sample was then briefly cooled on ice and centrifuged at 13,000 rpm in a micro-centrifuge. The following were then added to the DNA solution:5 µl Oligo Labeling Buffer, 1 µl Bovine Serum Albumin, 1 µl Klenow Fragment DNA Polymerase and all contents were mixed together and centrifuged at 13,000 rpm in a micro-centrifuge for 2 seconds. Finally 2.5 µl of ^{32}P-dCTP was added and the mixture was incubated for 4h at room temperature or 30 minutes at 37C to allow the labeling reaction to proceed.

Rapid DNA Precipitation: 1 µl of 0.5M EDTA (pH 8.0) was first added to chelate divalent cations. Two volumes of 5M ammonium acetate and 3 volumes of isopropyl alcohol were then added and the solution was mixed well. The sample was then incubate for 10 minutes at room temperature and then centrifuged at 13,000 rpm in a micro-centrifuge for 10 minutes.

Restriction Enzyme digests: All restriction digests were carried out in the suppliers recommended REact buffer system (Gibco) at the recommended optimum temperature for the enzyme.

RNase treatment: RNase treatments were carried out by the addition of 1µl of RNAse I, previously heat treated to denature DNAses, to each sample during restriction enzyme digests.

Southern Blot Analysis (Southern, 1975): Samples were first electrophoresed on an agarose gel for a suitable amount of time. Following electrophoreses the gel was transferred to a plastic dish and soak 2 x 15 minutes in Depurinating solution (0.25 M

HCl) and then washed with distilled water for 5 minutes. Next, the gel was treated for 30 minutes with Denaturing solution(0.5 M NaOH, 1.5 M NaCl) and then washed with distilled water for 5 minutes. Finally, the gel was soaked for 30 minutes in neutralising solution (solution (0.5 M Tris-HCl pH 7.4, 1.5 M NaCl) and then washed with distilled water for 5 minutes. The DNA fragments were then transferred from the gel onto a nylon membrane by capillary action (Southern Blotting) using 20x SSC as the transfer buffer for either 4 hours or overnight. Following the transfer the DNA was fixed onto the membrane by UV light. The membrane was washed for 15 minutes with Church-Gilbert Buffer and incubated at 65C for approximately 4h in order to block non-specific probe binding. The radioactively labeled probe was then added and the incubated for 4h or overnight in 10-20ml Church-Gilbert Buffer. Following incubation, the filter was washed 2-3 times at 65C with 3x SSC, 0.1% SDS and then autoradiographed.

6.3.2 E. coli DNA Techniques

Alkaline lysis mini-preps Mini-preps were carried out according to Maniatas *et al* (1982)

Preparation of Competent Cells and *E. coli* Transformation $CaCl_2$ competent cell preparation and transformations were carried out essentially according to Hanahan (1983)

Large Scale Plasmid Purification from *E. coli* (Qiagen): All large scale plasmid preparations were carried out using the Qiagen plasmid midi-preparation kit and performed according to the manufacturers instructions. In all large scale preparations, the *E. coli* strain was grown overnight at 37°C in a 2L flask containing 400 ml of Luria Broth supplemented with 50 μg\ml ampicillin.

6.3.3 Yeast DNA Techniques

Large Scale Linear Plasmid Preparation (Method 1)(Based on Reay, 1997): 400ml of the desired killer plasmid bearing strain of *K. lactis* were either grown overnight in YPD or for three days in SD supplemented with the appropriate amino acids at 30C. The cells were then harvested by centrifugation at 6,000 rpm for 10minutes at 4C and washed twice with 50ml distilled water. The pellet was resuspended in 25 ml Sorbitol EDTA buffer supplemented with Dithiothreitol and incubated for 30mins at 30C with mild agitation, harvested by centrifugation at 6,000 rpm for 10mins at 4C and washed with 50 ml 1.2M Sorbitol. The cell pellet was resuspended in 19ml SC and 5-10 mg Yeast Lytic Enzyme in SC were then added and protoplast formation was allowed to proceed at 30C, with its progress checked at 10 minute intervals by microscopy. When protoplast formation was approximately 90% complete the digestion was stopped by gentle centrifugation at 3,000 rpm for 5mins at 4C in a bench top centrifuge to remove the enzyme. The pellet was then washed twice in 25 ml 1.2M sorbitol and resuspended in 25ml 0.3M Running Buffer containing 1mM PMSF(75 µl) and 250 µl 10mg/ml RNase. A 5% NP-40 solution was added to a final concentration of 0.5%. The solution was gently stirred for 15 minutes on ice with a glass rod, which gently lysed the protoplasts. The cell debris was then removed by a high speed centrifugation, at 15,000 rpm for 15 minutes at 4C and the supernatant loaded onto a 20ml DE52 column. The column was washed with 0.3M NaCl RB, 0.4M NaCl RB and the linear plasmids eluted with 1.0M NaCl RB.

Large Scale Linear Plasmid Preparation (Method 2): Preparation of *K. lactis* cells was as above until the resuspension of the cell pellet in 25m 0.3M RB. Following resuspension of the protoplast pellet, solid ammonium sulphate was added to 20% of saturation and

incubated with mild agitation at 4C for 30 minutes. The solution was then centrifuged at 6000 rpm in a benchtop centrifuge for 30 minutes, the supernatant removed. Solid ammonium sulphate was added to a final concentration of 40% of saturation and the cycle was repeated. This was continued, in 20% of saturation steps until 80% of saturation. Following the 80% of saturation centrifugation a step of 10% of saturation to a final concentration of 90% of saturation was performed. Following incubation and centrifugation the supernatant was removed and the pellet retained. The pellet contained the purified k1 and k2 plasmids. The pellet was then resuspended in 0.3M RB and dialysed against 1 L of 0.3M RB overnight at 4C in the presence of 1mM PMSF.

Preparation of and transformation of competent *Kluyveromyces lactis*: Based on the *S. cerevisiae* transformation described by Gietz *et al* (1992). The culture was grown overnight in YEPD, subcultured into 50ml fresh medium (approximately 10^6 cells/ml maximum) and shaken at 30C until it reached a density of approximately 10^7/ml cells. Cells were harvested by centrifugation for 5 minutes at 5000 rpm and the cell pellet was resuspended in 10ml TE buffer. Following a second round of centrifugation, cells were resuspended in 5ml LA[0.1M LiAc in TE] buffer and incubated with shaking at 30C for 1 hour. The cells were then harvested by centrifugation and resuspended in 5 ml LAG [LA with 15% glycerol]. The cells were then dispensed into 300 µl aliquots and frozen on dry ice/IMS (Cells remained competent for 2 approximately months). For the actual transformation procedure, approximately 10 µl of the desired DNA species dissolved in TE (0.1 µg-10 µg) were then added to 0.3 ml competent cells, followed by 700 µl of PEG solution. [50 % PEG_{4000} warmed to dissolve and filtered in distilled water]. The mixture was then inverted 6-8 times and incubated at 30C for 1h. This incubation period was followed by a 5 minute "heat shock" at 42C. The cells were then harvested by

centrifugation and, resuspended in distilled water and plated on the desired selective media.

6.4 Protein Methods

Raising of antibodies in chickens As described in Gassman *et al* (1990). 20 μg of the GST fusion protein was emulsified with 1ml Freund's complete adjuvant and 500 μl was injected in each of the pectoral muscles of two chickens. A 10 μg booster injection was then carried out on day 12. Following the booster the chickens became ill and the third and final 10 μg booster was abandoned. On day 20, when the chickens resumed laying, eggs were harvested and the IgY component purified for analysis.

IgY Antibody and purification As described in Gassman *et al* (1990) Preparation: For four eggs, yolks were separated from egg whites, and briefly washed with dH_2O. The yolk skins were then removed, yielding approximately 60ml of yolk. 60 ml of Buffer A (10mM Potassium Phosphate pH 7.2) were then added to the yolks, followed by 120 ml Buffer A + 7% PEG6000 (3.5% PEG). This solution was then gently stirred at room temperature for 30 mins. Precipitate was removed by high speed centrifugation at 13,500 RPM for 10 mins at 4C and the supernatant was filtered through milk filter. Solid PEG 6000 was then added to a final concentration of 12% and the solution stirred until all PEG was dissolved. Precipitate was removed by centrifugation at 13,500 for 10 minutes at 4C. The supernatant was removed and the pellet resuspended in Buffer A. 40 ml Buffer of A + 24% PEG 6000 was then added. The mixture was then centrifuged at 13,500RPM at 4C for 10 minutes. The

supernatant was removed and the pellet resuspended in 20 ml Buffer A and the mixture was left to dialyze overnight against buffer A at 4C. The following day the dialysate was removed and centrifuged at 13,500 RPM for 10 minutes at 4C in order to remove

precipitated particles. The IgY solution was aliquoted into 1ml fractions into microcentrifuge tubes and stored at 4C.

Imobulin Band Purification As described in by Szewczyk and Summers (1992)

Small scale GST and GST-fusion purification (Based on Smith and Johnson, 1988) 1 ml samples of L Broth supplemented with ampicillin to a final concentration of 50 µg/ml were inoculated with a small loop-full of *E. coli* containing the appropriate plasmid and incubated overnight at 37°C with shaking. The samples were then diluted 1:10 the following day to a total volume of 5 ml with fresh L Broth containing 50 µg/ml ampicillin and grown for two hours. Cultures were then induced by the addition of 50 mg/ml IPTG to a final concentration of 1 mM and incubated for a further 3 hours. Aliquots (1.5 ml) were then centrifuged at high speed in a microcentrifuge for 5 seconds and re-suspended in 500 µl distilled water. These samples were then sonicated at for 5 x 5 seconds and cooled on ice between sonication bursts. Cell debris was removed by a high speed centrifugation and 10 µl of a 50% glutathione sepharose bead suspension was added to the supernatant. Samples were incubated 2 minutes during which they were constantly mixed by inversion. Beads were then spun down by centrifuging for 5 secs and supernatant removed. 20 µl of 2x SDS Sample buffer was then added and following boiling for 5 minutes, samples were ready loaded onto SDS-PAGE gels.

Large scale GST and GST-fusion protein purification (Based on Smith and Johnson, 1988) 5 ml samples of L Broth supplemented with ampicillin to a final concentration of 50 µg/ml were inoculated with a small loop-full of *E. coli* containing the appropriate plasmid and incubated overnight at 37°C with shaking. The samples were then diluted 1:100 the following day to a total volume of 400 ml with fresh L Broth containing 50 µg/ml ampicillin and grown for two hours. Cultures induced by the addition of 50 mg/ml IPTG to a final concentration of 1 mM and incubated for a further 3 hours. Bacteria were

harvested by centrifugation (5,000 RPM for 10 minutes in a Sorvall GS3 Rotor) and re-suspended in 20 ml distilled water. Cell debris was removed by high speed centrifugation. The supernatant was then poured through a 5ml column containing glutathione sepharose beads. The column was washed with 10 volumes of wash buffer (50 mM Tris (pH 7.5), 150 mM NaCl) and the protein eluted in 2 volumes of wash buffer containing 10 mM glutathione.

Large scale MBP and MBP-fusion purification (Based on Aitken *et al*, 1994) 5 ml samples of L Broth supplemented with ampicillin to a final concentration of 50 μg/ml were inoculated with a small loop-full of *E. coli* containing the appropriate plasmid and incubated overnight at 37°C with shaking. The samples were then diluted 1:10 the following day to a total volume of 400 ml with fresh L Broth containing 50 μg/ml ampicillin and grown for two hours. Cultures were then induced by the addition of 50 mg/ml IPTG to a final concentration of 1 mM and incubated a further 3 hours. Bacteria were harvested by centrifugation (5,000 RPM for 10 minutes in a Sorvall GS3 Rotor) and re-suspended in 20 ml distilled water. Cell debris was removed by high speed centrifugation. The supernatant was then poured through a 5ml column containing an amylose bead resin. The column was washed with 10 volumes of wash buffer (20 mM Tris (7.4), 0.2M NaCl, 10 mM β-mercaptoethanol) and the protein eluted in 2 volumes wash buffer containing 10 mM Maltose.

Polyacrylamide gels All polyacrylamide SDS-PAGE gels were made according to standard methods based on Laemmli (1970). Typically a 10% or 15% (v/v) separating and a 6.6% stacking gel were used. The acrylamide bis-acrylamide ratio was 37.5:1.

Silver Staining Silver staining was carried out essentially as described by Dunn and Crisp (1994).

Thrombin cleavage of GST and GST fusion Protein: Thrombin cleavage was usually performed by adding approximately 1-5µl of a 10µg/ml solution of thrombin to 20µl of a 450µg/ml sample of GST-fusion protein followed by overnight incubation at room temperature in Thrombin cleavage buffer.

Trichloroacetic Acid Protein (TCA) Precipitation: Originally based on Wright *et al.* (1988). 100% w/v TCA was added to the desired sample to a final concentration of 10% v/v. The sample was then incubated for 10 minutes on ice and centrifuged at 13,000 RPM in a micro-centrifuge for 10 minutes. The supernatant was then removed, the precipitate washed with one volume of ice cold ether and left to air dry for 10 minutes at room temperature.

Western Blotting: Gels were Electro-blotted using the Bio-Rad protean electro-blotting kit according to manufacturers instructions, onto nitrocellulose (Schleicher and Schuell, 0.45 µm. Following transfer, the membrane was treated with a Ponceau S-TCA stain for 10 minutes and the molecular weight marker bands marked with pencil. The membrane was then destained in P(T)BS. The membrane was then pre-blocked with 5% Marvel (non-fat dried milk). Next, the primary antibody (1:1000-1:10,000 dilution) was added and incubated for 4 hours or overnight. The membrane was then washed 4x 15 mins with 50 ml P(T)BS. Next the secondary antibody (1:10,000 dilution) was added and incubated for 1 hour. The membrane was then washed 4x 15 minutes with 50ml P(T)BS. For Illumin8 chemiluminescent Western detection the developing solution was made and added (10ml 100mM Tris (8.8), 100 µl Lumin 8A (100x Stock), 100 µl Lumin 8B (100x Stock) and 5 µl H_2O_2.

6.5 Other Methods

Dialysis Tubing preparation: The tubing was cut into appropriate lengths. The tubing was then immersed in 2% w/v Sodium bicarbonate and 1mM EDTA (pH 8) and autoclaved. The tubing was then immersed in distilled H_2O. The tubing was then soaked in 1mM EDTA for 10 mins. Finally, the tubing was cooled and stored in 50% EtOH/50% distilled water at 4C.

Long Term Storage of *E. coli* Strains: A 5ml culture of bacteria was grown overnight. 1ml of culture was then transferred to a fresh tube and 0.15 ml sterile glycerol added and the mixture vortexed before freezing for long term storage at -70C.

Long Term Storage of Yeast Strains: A 5ml culture of the desired yeast strain was grown overnight in minimal medium. The following day, an equal volume of yeast preservation mix (12.6 g/l K_2HPO_4, 3.6 g/l KH_2PO_4, 0.9 g/l trisodium citrate, 0.18 g/l $MgSO_4.7H_2O$, 1.8 g/l $(NH_4)_2SO_4$, 20% v/v glycerol) and the cells were quick frozen in a dry ice-IMS bath and samples were stored at -70C.

Chapter 7: Bibliography

Adhya, S., Shneidman, P., & Hurwitz, J. (1986). Reconstruction of adenovirus replication origins with a human nuclear factor-I binding-site. *JOURNAL OF BIOLOGICAL CHEMISTRY, **261**(7)*, 3339-3346.

Aitken, R., Gilchrist, J., Sinclair, M. (1994) Vectors to facilitate the creation of translational fusions to the maltose-binding protein of *Escherichia coli. GENE, **144**(1)*, 69-73

Almasan, A. & Mishra, N. (1988). Molecular characterization of the mitochondrial-DNA of a new stopper mutant ER-3 of *Neurospora crassa. GENETICS, **120**(4)*, 935-945

Ambrose, C. (1993). Replication of the linear plasmids of *Kluyveromyces lactis*: Analysis of the putative DNA polymerases. Ph.D. Thesis, University of Leicester, UK.

Amegadzie, B., Holmes, M. Cole, N., Jones, E., Moss, B.(1991). Identification, sequence and expression of the gene encoding the 2nd largest subunit of the Vaccinia virus DNA-dependent RNA-polymerase, *VIROLOGY **180**(1)*, 88-98

Ariga, H. (1983). Effect of aphidicolin on the elongation step of adenovirus DNA replication *in vitro. BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, **113**(1)*, 87-95.

Bamford, D., Rouhiainen, L., Takkinen, K., & Soderlund, H. (1981). Comparison of the lipid-containing bacteriophage-PRD1, bacteriophage- pr3, bacteriophage-pr4, bacteriophage-pr5 and bacteriophage-117. *JOURNAL OF GENERAL VIROLOGY, **57,** 365-373.*

Bamford, D., & Mindich, L. (1982). Structure of the lipid-containing bacteriophage-PRD1 - disruption of wild-type and nonsense mutant phage particles with guanidine- hydrochloride. *JOURNAL OF VIROLOGY, **44**(3)*, 1031-1038.

Bamford, D., Mcgraw, T., Mackenzie, G., & Mindich, L. (1983). Identification of a protein-bound to the termini of bacteriophage- PRD1 DNA. *JOURNAL OF VIROLOGY, **47**(2)*, 311-316.

Bamford, D., & Mindich, L. (1984). Characterization of the DNA protein complex at the termini of the bacteriophage PRD1 genome. *JOURNAL OF VIROLOGY, **50**(2)*, 309-315.

Bamford, J., & Bamford, D. (1990). Capsomer proteins of bacteriophage PRD1, a bacterial-virus with a membrane. *VIROLOGY, 177*(2), 445-451.

Bamford, J., & Bamford, D. (1991). Large-scale purification of membrane-containing bacteriophage PRD1 and its subviral particles. *VIROLOGY, 181*(1), 348-352.

Bamford, J., Hanninen, A., Pakula, T., Ojala, P., Kalkkinen, N., Frilander, M., & Bamford, D. (1991). Genome organization of membrane-containing bacteriophage-PRD1. *VIROLOGY, 183*(2), 658-676.

Bamford, J., Luo, C., Juuti, J., Olkkonen, V., & Bamford, D. (1993). Topology of the major capsid protein-p3 of bacteriophage PRD1 - analysis using monoclonal-antibodies and c-terminally truncated proteins. *VIROLOGY, 197*(2), 652-658.

Bamford, D., Caldentey, J., & Bamford, J. (1995). Bacteriophage PRD1 - a broad-host-range dsDNAtectivirus with an internal membrane. *ADVANCES IN VIRUS RESEARCH, 45*, 281-319.

Bernad, A., Zaballos, A., Salas, M. and Blanco, L. (1987). Structural and functional relationships between prokaryotic and eukaryotic DNA polymerases. *EMBO JOURNAL.* **6**, 4219-4225.

Bernad, A., Blanco, L. and Lazaro, J. M. (1989). A conserved 3' to 5' exonuclease active site in prokaryote and eukaryote DNA polymerases. *CELL* **59**, 219-228.

Bertrand, H., Chan,B., Griffiths, A. (1985) Insertion of a foreign nucleotide sequence into mitochondrial DNA causes senescence in *Neurospora intermedia. CELL, 41*(3), 877-884.

Betrand, H., Griffith, A. (1988) Abstracts of the 14th EMBO Annual Symposium, 7-8

Bischoff, D., Slavicek, J. (1996) Characterization of the *Lymantria dispar* nucleopolyhedrovirus 25K FP gene. *JOURNAL OF GENERAL VIROLOGY, 77*(8),1913-1923

Blanco, L., Bernadt, A., Esteban, J. A., & Salas, M. (1992). DNA-Independent Deoxynucleotidylation of the ϕ29 Terminal Protein by the ϕ29 DNA Polymerase. *THE JOURNAL OF BIOLOGICAL CHEMISTRY, 267*(2), 1225-1230.

Blanco, L., Lazaro, J., Devega, M., Bonnin, A., & Salas, M. (1994). Terminal protein-primed DNA amplification. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 91*(25), 12198-12202.

Blanco, L., & Salas, M. (1996). Relating structure to function in ϕ-29 DNA polymerase. *JOURNAL OF BIOLOGICAL CHEMISTRY, 271*(15), 8509-8512.

Blasco, M., Lazaro, J., Blanco, L., & Salas, M. (1993). Phi-29 DNA-polymerase active-site - residue asp(249) of conserved amino acid motif DX(2)SLYP is critical for synthetic activities. *JOURNAL OF BIOLOGICAL CHEMISTRY, **268***(32), 24106-24113.

Blasco, M., Lazaro, J., Blanco, L., & Salas, M. (1993). Phi-29 DNA-polymerase active-site - the conserved amino acid motif KX3NSXYG is involved in template-primer binding and dNTP selection. *JOURNAL OF BIOLOGICAL CHEMISTRY, **268***(22), 16763-16770.

Blasco, M., Mendez, J., Lazaro, J., Blanco, L., & Salas, M. (1995). Primer terminus stabilization at the ϕ-29 DNA polymerase active- site - mutational analysis of conserved motif KXY. *JOURNAL OF BIOLOGICAL CHEMISTRY, **270***(6), 2735-2740.

Boyd, V., Bozzini, M., Guga, P., Defranco, R., Yuan, P., Loudon, G., & Nguyen, D. (1995). Activation of the carboxy-terminus of a peptide for carboxy-terminal sequencing. *JOURNAL OF ORGANIC CHEMISTRY, **60***(8), 2581-2587.

Bravo, A., Hermoso, J., & Salas, M. (1994). *In vivo* functional-relationships among terminal proteins of *Bacillus subtilis* ϕ-29 related phages. *GENE, **148***(1), 107-112.

Brough, D., Droguett, G., Horwitz, M., & Klessig, D. (1993). Multiple functions of the adenovirus DNA-binding protein are required for efficient viral-DNA synthesis. *VIROLOGY, **196***(1), 269-281.

Broyles, S.S. and Moss, B. (1987a). Identification of the vaccinia virus gene encoding nucleoside triphosphate phosphohydrolase I, a DNA-dependent ATPase. *JOURNAL OF VIROLOGY* **61**, 1738-1742.

Bult C., White, O., Olsen, G., Zhou, L., Fleischmann, R., Sutton, G., Blank, J., Fitzerald, L., Clayton, R., Gocayne, J., Kerlavage, A., Dougherty, B., Tomb, J., Adams, M., Reich, C., Overbeek, R., Kirkness, E., Weinstock, K., Merrick, J., Glodek, A., Scott, J., Geoghagen, N., Weidman,J., Fuhrmann, J., Nguyen, D., Utterback, T., Kelley, J., Peterson, J., Sadow, P., Hanna, M., Cotton, M., Roberts, K., Hurst, M., Kaine, B., Borodovsky, M., Klenk, H., Fraser, C., Smith, H., Woese, C., Venter, J. (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii. SCIENCE, **273***(5278), 1058-1073.

Bussey, H., Sacks, W., Galley, D., & Saville, D. (1982). Yeast killer plasmid mutations affecting toxin secretion and activity and toxin immunity function. *MOLECULAR AND CELLULAR BIOLOGY, **2***(4), 346-354.

Butler, A., White, J., Folawiyo, Y., Edlin, A., Gardiner, D., & Start, M. (1994). 2 *Saccharomyces cerevisiae* genes which control sensitivity to G(1) arrest induced

by *Kluyveromyces lactis* toxin. *MOLECULAR AND CELLULAR BIOLOGY*, **14**(9), 6306-6316.

Butler, A., ODonnell, R., Martin, V., Gooday, G., & Stark, M. (1991). *Kluyveromyces lactis* Toxin has an essential chitinase activity. *EUROPEAN JOURNAL OF BIOCHEMISTRY*, **199**(2), 483-488.

Caldentey, J., Bamford, J., & Bamford, D. (1990). Structure and assembly of bacteriophage-PRD1, an *Escherichia coli* virus with a membrane. *JOURNAL OF STRUCTURAL BIOLOGY, **104***(1-3), 44-51.

Caldentey, J., Blanco, L., Savilahti, H., Bamford, D., & Salas, M. (1992). *In vitro* replication of bacteriophage PRD1 DNA- metal activation of protein-primed initiation and DNA elongation. *NUCLEIC ACIDS RESEARCH, **20***(15), 3971-3976.

Caldentey, J., Blanco, L., Bamford, D., & Salas, M. (1993). *In vitro* replication of bacteriophage PRD1 DNA - characterization of the protein-primed initiation site. *NUCLEIC ACIDS RESEARCH, **21***(16), 3725-3730.

Caravokyri, C., & Leppard, K. (1996). Human adenovirus type-5 variants with sequence alterations flanking the E2A gene - effects on E2 expression and DNA replication. *VIRUS GENES, **12***(1), 65-75.

Casini, G., & Revel, H. (1994). New small low-abundant nonstructural protein encoded by the L-segment of the dsRNA bacteriophage ϕ6. *VIROLOGY, **203***(2), 221-228.

Challberg, M., & Rawlins, D. (1984). Template requirements for the initiation of adenovirus DNA replication. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA-BIOLOGICAL SCIENCES, **81***(1), 100-104.

Challberg, M., Ostrove, J., & Kelly, T. (1982). Initiation of adenovirus DNA replication - detection of covalent complexes between nucleotide and the 80-kilodalton terminal protein. *JOURNAL OF VIROLOGY, **41***(1), 265-270.

Chan, R., Gonzalez, D. (1994) *PLANT PHYSIOLOGY. **106***. 1687-1688

Chan, B., Court, D., Vierula, P., & Bertrand, H. (1991). The *kalilo* linear senescence-inducing plasmid of *Neurospora* is an invertron and encodes DNA and RNA polymerases. *CURRENT GENETICS, **20***(3), 225-237.

Chang, J.(1985) Thrombin specificity - Requirement for apolar amino-acids adjacent to the thrombin cleavage site of polypeptide substrate. *EUROPEAN JOURNAL OF BIOCHEMISTRY. **1985***(151), 217-224

Chardonloriaux, I., Charpentier, M., & Percheron, F. (1986). Isolation and

characterization of a linear plasmid from *Streptomyces rimosus*. *FEMS MICROBIOLOGY LETTERS, 35*(2-3), 151-155.

Chen, X., Wesolowskilouvel, M., Tanguyrougeau, C., & Fukuhara, H. (1991). Promoter activity associated with the left inverted terminal repeat of the killer plasmid-k1 from yeast. *BIOCHIMIE, 73*(9), 1195-1203.

Chen, H., Ramachandra, M., & Padmanabhan, R. (1994). Biochemical-characterization of a temperature-sensitive adenovirus DNA-polymerase. *VIROLOGY, 205*(1), 364-370.

Church, G., Gilbert, W. (1984) Genomic sequencing. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA-BIOLOGICAL SCIENCES. 81*(7), 1991-1995

Ciufo, D., Mullen, M., & Hayward, G. (1994). Identification of a dimerization domain in the C-terminal segment of the IE110 transactivator protein from *Herpes simplex* virus. *JOURNAL OF VIROLOGY, 68*(5), 3267-3282.

Clark-Walker, G., Miklos, G. (1974) Localization and quantitation of circular DNA in yeast. *EUROPEAN JOURNAL OF BIOCHEMISTRY. 41*, 359-365.

Coenjaerts, F., & Vandervliet, P. (1994). Early dissociation of nuclear factor-I from the origin during initiation of adenovirus DNA replication studied by origin immobilization. *NUCLEIC ACIDS RESEARCH, 22*(24), 5235-5240.

Coenjaerts, F., Vanoosterhout, J., & Vandervliet, P. (1994). The OCT-1 POU domain stimulates adenovirus DNA-replication by a direct interaction between the viral precursor terminal protein-DNA polymerase complex and the POU homeodomain. *EMBO JOURNAL, 13*(22), 5401-5409.

Coenjaerts, F., & Vandervliet, P. (1995). Adenovirus DNA replication in a reconstituted system. *METHODS IN ENZYMOLOGY, 262*, 548-560.

Cong, Y., Wesolowskilouvel, M., & Fukuhara, H. (1994). Creation of a functional promoter by rearrangement in a *Kluyveromyces lactis* linear plasmid. *GENE, 147*(1), 125-129.

Cong, Y., Yarrow, D., Li, Y., & Fukuhara, H. (1994). Linear DNA plasmids from *Pichia etchellsii, Debaryomyces hansenii* and *Wingea robertsiae*. *MICROBIOLOGY-UK, 140*(Pt6), 1327-1335.

Conley, E., Saunders, V., & Saunders, J. (1986). Deletion and rearrangement of plasmid DNA during transformation of *Escherichia coli* with linear plasmid molecules. *NUCLEIC ACIDS RESEARCH, 14*(22), 8905-8917.

Conner, J., Cross, A., Murray, J., & Marsden, H. (1994). Identification of structural

domains within the large subunit of *Herpes simplex* virus ribonucleotide reductase. *JOURNAL OF GENERAL VIROLOGY, 75*(Pt12), 3327-3335.

Court, D., Griffiths, A., Kraus, S., Russell, P., & Bertrand, H. (1991). A new senescence-inducing mitochondrial linear plasmid in field- isolated *Neurospora crassa* strains from India. *CURRENT GENETICS, 19*(2), 129-137.

Cowman, A., Bernad, O., Stewart, N., Kemp, D. (1984) Genes of the protozoan parasite *Babesia bovis* that rearrange to produce RNA species with different sequences. *CELL, 37*(2), 653-660

Dekker, J., Vanoosterhout, J., & Vandervliet, P. (1996). 2 regions within the DNA binding domain of Nuclear Factor I interact with DNA and stimulate adenovirus DNA replication independently. *MOLECULAR AND CELLULAR BIOLOGY, 16*(8), 4073-4080.

Dekker, J., Kanellopoulos, P., Loonstra, A., Vanoosterhout, J., Leonard, K., Tucker, P., & Vandervliet, P. (1997). Multimerization of the adenovirus DNA binding protein is the driving force for ATP-independent DNA unwinding during strand displacement synthesis. *EMBO JOURNAL, 16* (6), 1455-1463.

Delouvencourt, L., Fukuhara, H., Heslot, H., & Wesolowski, M. (1983). Transformation of *Kluyveromyces lactis* by killer plasmid DNA. *JOURNAL OF BACTERIOLOGY, 154*(2), 737-742.

Devries, E., Vandriel, W., Tromp, M., Vanboom, J., & Vandervliet, P. (1985). Adenovirus DNA replication *in vitro* - site-directed mutagenesis of the nuclear factor-I binding-site of the *ad2* origin. *NUCLEIC ACIDS RESEARCH, 13*(13), 4935-4952.

Diffley, J., & Stillman, B. (1986). Purification of a cellular, double-stranded DNA binding protein required for initiation of adenovirus DNA replication by using a rapid filter-binding assay. *MOLECULAR AND CELLULAR BIOLOGY, 6*(5), 1363-1373.

Dobbs, L., Zhao, L., Sripad, G., & Padmanabhan, R. (1990). Mutational analysis of single-stranded-DNA templates active in the *in vitro* initiation assay for adenovirus DNA replication. *VIROLOGY, 178*(1), 43-51.

Dunn, M., Crisp, S. (1994) Detection of proteins in polyacrylamide gels using an ultrasensitive silver staining technique. In: Methods in Molecular Biology, *32*: Basic Protein and Peptide Protocols.

Enomoto, T., Lichy, J., Ikeda, J., & Hurwitz, J. (1981). Adenovirus DNA replication *in vitro* - purification of the terminal protein in a functional form. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA-BIOLOGICAL SCIENCES, **78***(11), 6779-6783.

Esser, K., Kempken, F. (1986) Structure and function of linear extrachromosmal DNA in Eukaryotes. *PROCESS BIOCHEMISTRY, **21***(3). 69-76

Esteban, J., Soengas, M., Salas, M., & Blanco, L. (1994). 3'->5' exonuclease active-site of ϕ-29 DNA polymerase - evidence favoring a metal ion-assisted reaction-mechanism. *JOURNAL OF BIOLOGICAL CHEMISTRY, **269***(50), 31946-31954.

Esteban, J., Blanco, L., Villar, L., & Salas, M. (1997). *In vitro* evolution of terminal protein-containing genomes. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, **94***(7), 2921-2926.

Foster, G., Ackrill, A., Goldin, R., Kerr, I., Thomas, H., & Stark, G. (1995). Expression of the terminal protein region of hepatitis B virus inhibits cellular-responses to interferon-α and interferon-γ and double-stranded-RNA *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, **92***(8), 3632-3632.

Francou, F., Randscholt, N., Decaris, B., & Gregoire, A. (1987). Presence of several plasmids in the fungus *Ascobolus immersus*, and homologies with the linear plasmid pA1. *JOURNAL OF GENERAL MICROBIOLOGY, **133***(FEB), 311-316.

Fraser, C., Gocayne, J., White, O., Adams, M., Clayton, R., Fleischman, R., Bult, C., Kerlavage, A., Sutton, G., Kelley, J., Fritchman, J., Weidman, J., Small, K., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T., Saudek, D., Phillips, C., Merrick, J., Tomb, J., Dougherty, B., Bott, K., Hu, P., Lucier, T., Peterson, S., Smith, H., Hutchison, C., Venter, J. (1995) The minimal gene complement of *Mycoplasma genitalium. SCIENCE, **270***(5235), 397-403

Freije, J., Lain, S., Vinuela, E., & Lopezotin, C. (1993). Nucleotide-sequence of a nucleoside triphosphate phosphohydrolase gene from African Swine Fever Virus. *VIRUS RESEARCH, **30***(1), 63-72.

Freire, R., Serrano, M., Salas, M., & Hermoso, J. (1996). Activation of replication origins in ϕ-29-related phages requires the recognition of initiation proteins to specific nucleoprotein complexes. *JOURNAL OF BIOLOGICAL CHEMISTRY, **271***(48), 31000-31007.

Fukuda, K., Maebuhi, M., Takata, H., & Gunge, N. (1997). The Linear Plasmid pDHL1 from *Debaryomyces hanseii* Encodes a Protein Highly Homologous to the pGKL1-Plasmid DNA Polymerase. *YEAST, **13***, 613-620.

Futterer, J., & Winnacker, E. (1984). Adenovirus DNA replication. *CURRENT TOPICS IN MICROBIOLOGY AND IMMUNOLOGY, **111***, 41-64.

Gallo, E. and Galeotti, C. L. (1990). Transplacement of a nuclear gene into the cytoplasmic killer plasmids of *Kluyveromyces lactis*. Abstracts of the XV International Conference of Yeast Genetics and Molecular Biology, John Wiley & Sons, Inc., New York, London, p. 547.

Gamas, P., Niebel, F., Lescure, N., Cullimore, J. (1996) Use of a subtractive hybridization approach to identify new *Mediago truncatula* genes induced during root nodule development. *MOLECULAR PLANT-MICROBE INTERACTIONS **9***(4), 233-242

Gassman, M., Thommes, P., Weiser, T., Hubscher, U. (1990) Efficient production of chicken egg yolk antibodies against a conserved mammalian protein. *FASEB JOURNAL. **4***. 2528-2532

Gerendasy, D., & Ito, J. (1987). The genome of lipid-containing bacteriophage-PRD1, which infects gram-negative bacteria, contains long, inverted terminal repeats. *JOURNAL OF VIROLOGY, **61***(2), 594-596.

Gerendasy, D., & Ito, J. (1990). Nucleotide-sequence and transcription of the right early region of bacteriophage PRD1. *JOURNAL OF BACTERIOLOGY, **172***(4), 1889-1898.

Giasson, L., & Lalonde, M. (1987). Analysis of a linear plasmid isolated from the pathogenic fungus. *Ceratocystis fimbriata ell* and *halst. CURRENT GENETICS, **11***(4), 331-334.

Giraudpanis, M., & Lilley, D. (1996). T4 endonuclease VII - importance of a histidine-aspartate cluster within the zinc-binding domain. *JOURNAL OF BIOLOGICAL CHEMISTRY, **271***(51), 33148-33155.

Gounari, F., Defrancesco, R., Schmitt, J., Vandervliet, P., Cortese, R., & Stunnenberg, H. (1990). Amino terminal domain of NF1 binds to DNA as a dimer and activates adenovirus DNA replication. *EMBO JOURNAL, **9***(2), 559-566.

Graham, F. (1996). Adenovirus vectors for gene-expression and gene-therapy. *TRANSFUSION SCIENCE, **17***(1), 15-27.

Grahn, A., Bamford, J., Oneill, M., & Bamford, D. (1994). Functional-organization of the bacteriophage PRD1 genome. *JOURNAL OF BACTERIOLOGY, **176***(10), 3062-3068.

Grimes, S., & Anderson, D. (1997). The bacteriophage ϕ29 packaging proteins supercoil the DNA ends. *JOURNAL OF MOLECULAR BIOLOGY, **266***(5), 901-914.

Guerineau, M., Grandchamp, C., Paoletti, C., Slonimski, P. (1971) Characterization of a new class of circular DNA molecules in yeast. *BIOCHEMICAL BIOPHYSICAL RESEARCH COMMUNITY. **42**.* 550-557

Gunge, N., Tamaru, A., Ozawa, F., & Sakaguchi, K. (1981). Isolation and Characterization of Linear Deoxyribonucleic Acid Plasmids from *Kluyveromyces lactis* and the Plasmid Associated Killer Character. *JOURNAL OF BACTERIOLOGY, **145***(1), 382-390.

Gunge, N. and Yamane, C. (1985). Incompatibility of linear DNA killer plasmids, pGKL1 and pGKL2, from *Kluyveromyces lactis* with mitochondrial DNA from *Saccharomyces cerevisiae. JOURNAL OF BACTERIOLOGY* **159**, 533-539.

Gunge, N., Fukuda, K., Morikawa, S., Murakami, K., Takeda, M., & Miwa A. (1993). Osmophilic linear plasmids from the salt-tolerant *Debaryomyces hansenii. CURRENT GENETICS, **23***, 443-449.

Gunge, N., Takahashi, S., Fukuda, K., Ohnishi, T., & Meinhardt, F. (1994). UV hypersensitivity of yeast linear plasmids. *CURRENT GENETICS, **26***(4), 369-373.

Gunge, N., Fukuda, K., Takahashi, S., & Meinhardt, F. (1995). Migration of the yeast linear DNA plasmid from the cytoplasm into the nucleus in *Saccharomyces cerevisiae. CURRENT GENETICS, **28***(3), 280-288.

Hanninen, A., Bamford, D., & Bamford, J. (1997). Assembly of membrane-containing bacteriophage PRD1 is dependent on *groEL* and *groES. VIROLOGY, **227***(1), 207-210.

Hashiba, T., Homma, Y., Hyakumachi, M., Matsuda, I. (1984) Isolation of a DNA plasmid in the fungus *Rhizoctonia solani. JOURNAL OF GENERAL MICROBIOLOGY **130***, 2067-2070

Hashiba, T. (1990). Linear plasmid DNAs of the plant pathogenetic fungus. *NIPPON NOGEIKAGAKU KAISHI-JOURNAL OF THE JAPAN SOCIETY FOR BIOSCIENCE BIOTECHNOLOGY AND AGROCHEMISTRY, **64***(1), 58-61.

Hatfield, L., & Hearing, P. (1993). The NFIII/OCT-1 binding-site stimulates adenovirus DNA replication *in vivo* and is functionally redundant with adjacent sequences. *JOURNAL OF VIROLOGY, **67***(7), 3931-3939.

Hay, R. (1985). Origin of adenovirus DNA replication - role of the Nuclear Factor I binding-site *in vivo. JOURNAL OF MOLECULAR BIOLOGY, **186***(1), 129-136.

Hay, R. (1985). The origin of adenovirus DNA replication - minimal DNA sequence requirement *in vivo. EMBO JOURNAL, **4***(2), 421-426.

Hay, R. (1986). Role of Nuclear Factor I in adenovirus DNA replication. *JOURNAL OF CELLULAR BIOCHEMISTRY,* **S10B**, 157-157.

Hay, R., Barrett, P., Clark, L., & Harris, M. (1987). The role of sequence-specific DNA binding proteins in adenovirus DNA replication. *JOURNAL OF CELL SCIENCE,* **S7**, 51-65.

Hay, R., Freeman, A., Leith, I., Monaghan, A., & Webster, A. (1994). Protein-DNA contacts and DNA structural transitions during initiation of adenovirus DNA replication. *JOURNAL OF CELLULAR BIOCHEMISTRY,* **S18C SIC**, 151-151.

Hayman, G., & Bolen, P. (1991). Linear DNA plasmids of *Pichia inositovora* are associated with a novel killer toxin activity. *CURRENT GENETICS, **19***(5), 389-393.

Hermanns, J. and Osiewacz, H.D. (1992). The linear mitochondrial plasmid pAL2-1 of a long lived *Podospora anserina* mutant is an invertron encoding a DNA and RNA polymerase. *CURRENT GENETICS* **22**, 491-500.

Hiratsuka, K., Namba, S., Yamashita, S., & Tsuchizaki, T. (1989). Linear plasmid DNA in *Pyrenophora graminea*, a fungal pathogen of barley. *PROCEEDINGS OF THE JAPAN ACADEMY SERIES B-PHYSICAL AND BIOLOGICAL SCIENCES, **65***(3), 42-44.

Hishinuma, F., Nakamura, K., Hirai, K., Nishiwaza, R., Gunge, N. and Maeda, T. (1984). Cloning and nucleotide sequence of the DNA killer plasmids from yeast. *NUCLEIC ACIDS RESEARCH.* **12**, 7581-7597.

Hishinuma, F., & Hirai, K. (1991). Genome organization of the linear plasmid, pSKL, isolated from *Saccharomyces kluyveri. MOLECULAR & GENERAL GENETICS, **226***(1-2), 97-106.

Hopkinson, S., Pollenz, R., Drummon, I., Chisholm, R. (1989) Expression and organisation of BP74, a cyclic AMP-regulated gene expressed during *Dictyostelium discoideum* development. *MOLECULAR AND CELLULAR BIOLOGY. **9***(10). 4170-4178

Hotta, J., Shi, L., & Ginsberg, H. (1994). Effect of *cd4* gene-expression on adenovirus replication. *JOURNAL OF VIROLOGY, **68*** (11), 7284-7291.

Hsieh, J., Jung, G., Leavitt, M., & Ito, J. (1987). Primary structure of the DNA terminal protein of bacteriophage PRD1. *NUCLEIC ACIDS RESEARCH, 15*(21), 8999-9009.

Hu, J., & Seeger, C. (1996). Expression and characterization of hepadnavirus reverse transcriptases. *METHODS IN ENZYMOLOGY, **275***, 195-208.

Hu, J., Toft, D. O., & Seeger, C. (1997). Hepadnavirus assembly and reverse transcription require a multi-component chaperone complex which is incorporated into nucleocapsids. *The EMBO Journal, **16***(1), 59-68.

Hussy, P., Langen, H., Mous, J., & Jacobsen, H. (1996). Hepatitis C virus core protein - carboxy-terminal boundaries of 2 processed species suggest cleavage by a signal peptide peptidase. *VIROLOGY, **224***(1), 93-104.

Illana, B., Blanco, L., & Salas, M. (1996). Functional-characterization of the genes-coding for the terminal protein and DNA polymerase from bacteriophage ga1 - evidence for a sliding-back mechanism during protein-primed ga1 DNA replication. *JOURNAL OF MOLECULAR BIOLOGY, **264***(3), 453-464.

Jung, G., Leavitt, M., & Ito, J. (1987). Yeast killer plasmid pGKL1 encodes a DNA polymerase belonging to the family B-DNA polymerases. *NUCLEIC ACIDS RESEARCH, **15***(21), 9088-9088.

Jung, G., Leavitt, M., Hsieh, J., & Ito, J. (1987). Bacteriophage PRD1 DNA polymerase - evolution of DNA polymerases. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, **84***(23), 8287-8291.

Jung, G., Leavitt, M., Schultz, M., & Ito, J. (1990). Site-specific mutagenesis of PRD1 DNA polymerase - mutations in highly conserved regions of the family-B DNA polymerase. *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, **170***(3), 1294-1300.

Kämper, J., Meinhardt, F., Gunge, N. and Esser, K. (1989a). *In vivo* construction of linear vectors based on killer plasmids from *Kluyveromyces lactis*: selection of a nuclear gene results in attachment of telomeres. *MOLECULAR AND CELLULAR BIOLOGY.* **9**, 3931-3937.

Kämper, J., Esser, K., Gunge, N. and Meonhardt, F. (1989b). New recombinant linear DNA elements derived from *Kluyveromyces lactis* killer plasmids. *NUCLEIC ACIDS RESEARCH.* **17**, 1781.

Kämper, J., Esser, K., Gunge, N., & Meinhardt, F. (1991). Heterologous gene-expression on the linear DNA killer plasmid from *Kluyveromyces lactis*. *CURRENT GENETICS, **19***(2), 109-118.

Kanegae, Y., Miyake, S., Sato, Y., Lee, G., & Saito, I. (1996). Adenovirus vector technology - an efficient method for constructing recombinant adenovirus and on/off switching of gene-expression. *ACTA PAEDIATRICA JAPONICA, **38***(2), 182-188.

Kempken, F., Meinhardt, F. and Esser, K. (1989). *In organello* replication and viral affinity of linear extrachromosomal DNA of the ascomycete *Ascobulus immersus*. *MOLECULAR AND GENERAL GENETICS* **218**, 523-530.

Kempken, F. (1989). A transfer RNA like structure from a linear plasmid of the fungus *Ascobolus immersus*. *NUCLEIC ACIDS RESEARCH, 17*(16), 6740-6740.

Kempken, F., Hermanns, J. and Osiewacz. H. D. (1992). Evolution of linear plasmids. *JOURNAL OF MOLECULAR EVOLUTION.* **35**, 502-513.

Kenny, M., & Hurwitz, J. (1988). Initiation of adenovirus DNA replication 2: Structural requirements using synthetic oligonucleotide adenovirus templates. *JOURNAL OF BIOLOGICAL CHEMISTRY, 263*(20), 9809-9817.

Kenny, M., Balogh, L., & Hurwitz, J. (1988). Initiation of adenovirus DNA replication 1: Mechanism of action of a host protein required for replication of adenovirus DNA templates devoid of the terminal protein. *JOURNAL OF BIOLOGICAL CHEMISTRY, 263*(20), 9801-9808.

Kikuchi, Y., Hirai, K., & Hishinuma, F. (1984). The yeast linear DNA killer plasmids, pGKL1 and pGKL2 possess terminally attached proteins. *NUCLEIC ACIDS RESEARCH, 12*(14).

Kim, W., Mcnabb, S., & Klassen, G. (1988). A linear plasmid in *Tilletia controversa*, a fungal pathogen of wheat. *CANADIAN JOURNAL OF BOTANY-JOURNAL CANADIEN DE BOTANIQUE, 66*(6), 1098-1100.

Kinashi, H., & Shimajimurayama, M. (1991). Physical characterization of scp1, a giant linear plasmid from *Streptomyces coelicolor*. *JOURNAL OF BACTERIOLOGY, 173*(4), 1523-1529.

King, A., & Vandervliet, P. (1994). A precursor terminal protein trinucleotide intermediate during initiation of adenovirus DNA replication - regeneration of molecular ends *in vitro* by a jumping back mechanism. *EMBO JOURNAL, 13*(23), 5786-5792.

King, A., Teertstra, W., Blanco, L., Salas, M., & Vandervliet, P. (1997). Processive proofreading by the adenovirus DNA polymerase. Association with the priming protein reduces exonucleolytic degradation. *NUCLEIC ACIDS RESEARCH, 25*(9), 1745-1752.

Kishi, T., Miura, K., Matsumoto, K., & Hirokawa, H. (1993). Complementation assay of primer protein - gene-expression systems of plasmid vectors support the infection of suppressor sensitive mutant phages M2 and ϕ-29. *JAPANESE JOURNAL OF GENETICS, 68*(4), 243-255.

Kishi, T., Miura, K., Matsumoto, K., & Hirokawa, H. (1994). Effects of altered RGD (arg-gly-asp) sequences in the primer protein on the protein-primed DNA-replication of bacteriophage-M2 *in vitro*. *JOURNAL OF GENERAL AND APPLIED MICROBIOLOGY, **40***(3), 197-208.

Kistler, H., & Leong, S. (1986). Linear plasmid-like DNA in the plant pathogenic fungus *Fusarium oxysporum* f sp *conglutinans*. *JOURNAL OF BACTERIOLOGY, **167***(2), 587-593.

Kitabayashi, I., Matsumoto, K., & Hirokawa, H. (1993). The primer protein of bacteriophage-M2 inhibits elongation activity of its DNA-polymerase. *JOURNAL OF GENERAL AND APPLIED MICROBIOLOGY, **39***(5), 467-478.

Kitada, K. and Hishinuma, F. (1987). A new linear DNA plasmid isolated from the yeast *Saccharomyces kluyveri*. *MOLECULAR & GENERAL GENETICS*, **206**, 377-381.

Kitada, K., & Gunge, N. (1988). Palindrome-hairpin linear plasmids possessing only a part of the ORF1 gene of the yeast killer plasmid pGKL1. *MOLECULAR & GENERAL GENETICS, **215***(1), 46-52.

Kozak, M. (1983) Translation of insulin-related polypeptides from messenger-RNAs with tandemly reiterated copies of the ribosome binding site. *CELL, **34***(3) 971-978

Laemmeli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *NATURE* **227**, 680-685.

Lally, C., Antoine, G., Dorper, T., Groger, W., & Winnacker, E. (1983). Initiation of the adenovirus DNA replication. *ZENTRALBLATT FÜR BAKTERIOLOGIE MIKROBIOLOGIE UND HYGIENE SERIES A- MEDICAL MICROBIOLOGY INFECTIOUS DISEASES VIROLOGY PARASITOLOGY, **255***(2-3), 176-176.

Lanford, R., Notvall, L., Lee, H., & Beames, B. (1997). Transcomplementation of nucleotide priming and reverse transcription between independently expressed tp and RT domains of the hepatitis B virus reverse transcriptase. *JOURNAL OF VIROLOGY, **71***(4), 2996-3004.

Larsen, M., Gunge, N., Meinhardt, F. (1998) *Kluyveromyces lactis* killer plasmid pGKL2: Evidence for a viral-like capping enzyme encoded by ORF3. *PLASMID. **40***(3), 243-246.

Lazaro, J., Blanco, L., & Salas, M. (1995). Purification of bacteriophage ϕ-29 DNA polymerase. *METHODS IN ENZYMOLOGY, **262***, 42-49.

Lee, Y., & Wickner, R. (1992). *Mak10*, a glucose-repressible gene necessary for replication of a dsRNA virus of *Saccharomyces cerevisiae*, has T-cell receptor alpha subunit motifs. *GENETICS, **132***(1), 87-96.

Leegwater, P., Rombouts, R., & Vandervliet, P. (1988). Adenovirus DNA replication *in vitro* - duplication of single-stranded-DNA containing a panhandle structure. *BIOCHIMICA ET BIOPHYSICA ACTA, **951***(2-3), 403-410.

Leegwater, P., Vandriel, W., & Vandervliet, P. (1985). Recognition site of Nuclear Factor I, a sequence-specific DNA binding protein from HeLa-cells that stimulates adenovirus DNA replication. *EMBO JOURNAL, **4***(6), 1515-1521.

Leibowitz, M., Thiele, D., Hannig, E. (1983) Structure of separated strands of double-stranded RNA from the killer virus of yeast: a translational model. In Compans, R.W. and and Bishop, D.H.L (Eds) *Double-stranded RNA Viruses*. Elsevier, Amsterdam, 457-462

Leroux, A., Kerdiles, B., Walls, D., Dedieu, J., & Perricaudet, M. (1994). The Epstein-Barr virus determined nuclear antigens EBNA-3A, EBNA-3B, and EBNA-3C repress EBNA-2-mediated transactivation of the viral terminal protein-1 gene promoter. *VIROLOGY, **205***(2), 596-602.

Lichy, J., Field, J., Horwitz, M., & Hurwitz, J. (1982). Separation of the adenovirus terminal protein-precursor from its associated DNA polymerase - role of both proteins in the initiation of adenovirus DNA replication. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA-BIOLOGICAL SCIENCES, **79***(17), 5225-5229.

Ligon, J., Bolen, P., Hill, D., Bothast, R., Kurtzman, C. (1989) Physical and biological characterization of linear DNA plasmids of the yeast *Pichia inositovora*. *PLASMID. **21***(3), 185-194

Lonergan, K., Gray, M. (1993) Predicted editing of additional transfer-RNAs in *Acanthamoeba castellanii* mitochondria. *NUCLEIC ACIDS RESEARCH. **21***(18), 4402

Malke, H., Noack, D., & Holm, S. (1982). Transformation of *Streptococcus sanguis (challis)* by linear plasmid molecules. *MOLECULAR & GENERAL GENETICS, **185***(3), 502-505.

Martin, A., Lopez, R., & Garcia, P. (1995). Nucleotide-sequence and transcription of the left early region of *Streptococcus pneumoniae* bacteriophage CP-1 coding for the terminal protein and the DNA polymerase. *VIROLOGY, **211***(1), 21-32.

Martin, A., Blanco, L., Garcia, P., Salas, M., & Mendez, J. (1996). *In vitro* protein-primed initiation of pneumococcal phage CP-1 DNA replication occurs at the 3rd 3'-nucleotide of the linear template - a stepwise sliding-back mechanism. *JOURNAL OF MOLECULAR BIOLOGY, **260***(3), 369-377.

Martin, A., Lopez, R., & Garcia, P. (1996). Analysis of the complete nucleotide-sequence and functional- organization of the genome of *Streptococcus pneumoniae* bacteriophage- CP-1. *JOURNAL OF VIROLOGY, 70*(6), 3678-3687.

Massung, R., Jayarama, V., Moyer, R. (1993) DNA-sequence analysis of conserved and unique regions of Swinepox virus - Identification of genetic elements supporting phenotypic observations including a novel G-protein coupled receptor homolog. *VIROLOGY. 197*. 511-528

Matsumoto K., Takano, H., Kim, C., Hirokawa, H. (1989) Primary structure of bacteriophage M2 DNA polymerase: conserved segments within protein-priming DNA polymerases and DNA polymerase I of *Escherichia coli. GENE, 84* 247-255

Mcgraw, T., Yang, H., & Mindich, L. (1983). Establishment of a physical and genetic-map for bacteriophage PRD1. *MOLECULAR & GENERAL GENETICS, 190*(2), 237-244.

Mcneel, D., & Tamanoi, F. (1991). Terminal region recognition factor I, a DNA binding protein recognizing the inverted terminal repeats of the pGKL linear DNA plasmids. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 88*(24), 11398-11402.

Meinhardt, F. and Esser, K. (1984). Linear extrachromosomal DNA in the morel *Morchella conica. CURRENT GENETICS.* **8**, 15-18.

Meinhardt, F., Kempken, F., & Esser, K. (1986). Proteins are attached to the ends of a linear plasmid in the filamentous fungus *Ascobolus immersus. CURRENT GENETICS, 11*(3), 243-246.

Meinhardt, F., Kempken, F. and Esser, K. (1986). Proteins are attached to the ends of a linear plasmid in the filamentous fungus *Ascobulus immersus. CURRENT GENETICS.* **11**, 243-246.

Meinhardt, F. and Esser, K. (1987). The plasmids of the morels: characterization and prerequisites for vector development. *APPLIED MICROBIOLOGY AND BIOTECHNOLOGY* **27**, 276-282.

Meinhardt, F., Kempken, F., Kämper, J. and Esser, K. (1990). Linear plasmids among eukaryotes: fundamentals and application. *CURRENT GENETICS.* **17**, 89-95.

Meinhardt, F. and Rohe, M. (1993). Extranuclear inheritance: Linear protein-primed replicating genomes in plants and microorganisms. In *Progress in Botany.* Behnke, H.D., Luettge, U., Esser, K., Kadereit, J.W. and Runge, M. (eds). Springer Verlag, Berlin/Heidelberg/New York. **54**, pp 334-357.

Meinhardt, F., Wodara, C., Larsen, M. and Schickel, J. (1994). A novel approach to expresss a heterologous gene on *Kluyveromyces lactis* linear killer plasmids: Expression of the bacterial *aph* gene from a cytoplasmic promoter fragment without in-phase fusion to the plasmid open reading frame. *PLASMID* **32**, 318-327.

Meinhardt, F. and Holtwick, R. (1995). Generation of a stable non-reverting Leu$^-$ mutant of *Kluyveromyces lactis* by gene disruption. *APPLIED MICROBIOLOGY AND BIOTECHNOLOGY.* **42**, 734-737.

Mendez, J., Blanco, L., Lazaro, J., & Salas, M. (1994). Primer-terminus stabilization at the ϕ29 dna-polymerase active- site - mutational analysis of conserved motif TX(2)GR. *JOURNAL OF BIOLOGICAL CHEMISTRY, **269***(47), 30030-30038.

Miyashita, S., Hirochika, H., Ikeda, J., & Hashiba, T. (1990). Linear plasmid DNAs of the plant pathogenic fungus *Rhizoctonia solani* with unique terminal structures. *MOLECULAR & GENERAL GENETICS, **220***(2), 165-171.

Mohan, M., Meyer, R., Anderson, J., Horgen, P. (1984) Plasmid-like DNAs in the commercially important mushroom genus *Agaricus*. *CURRENT GENETICS. **8***(8), 615-619.

Mul, Y., & Vandervliet, P. (1993). The adenovirus DNA binding protein effects the kinetics of DNA replication by a mechanism distinct from NFI or OCT-1. *NUCLEIC ACIDS RESEARCH, **21***(3), 641-647.

Nagata, K., Guggenheimer, R., & Hurwitz, J. (1983). Adenovirus DNA replication *in vitro* - synthesis of full-length DNA with purified proteins. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA-BIOLOGICAL SCIENCES, **80***(14), 4266-4270.

Nesterova, G., Semykina, L., & Filatov, A. (1981). A study of the killer plasmid mutants isolated by treatment with 5- fluorouracil. *GENETIKA, **17***(3), 389-398.

Noelling, J., Devos, W. (1992) Identification of the CTAG-recognizing restriction-modification systems MTHZI from *Methanobacterium thermoformicicum* and characterization of the plasmid encoded MTHZIM gene. *NUCLEIC ACIDS RESEARCH. **20***(19), 5047-5052

Normand, P., Simonet, P., Giasson, L., Ravelchapius, P., Fortin, J., & Lalonde, M. (1987). Presence of a linear plasmid-like DNA molecule in the fungal pathogen *Ceratocystis fimbriata ell* and *halst. CURRENT GENETICS, **11***(4), 335-338.

Oberhaus, S., & Newbold, J. (1993). Detection of DNA-polymerase activities associated with purified duck hepatitis B virus core particles by using an activity gel assay. *JOURNAL OF VIROLOGY, **67***(11), 6558-6566.

Oberhaus, S., & Newbold, J. (1995). Detection of an RNAse-H activity associated with hepadnaviruses. *JOURNAL OF VIROLOGY, **69***(9), 5697-5704.

Oberhaus, S., & Newbold, J. (1996). *In situ* DNA polymerase and RNAse-H activity gel assays as applied to hepadnavirus particles. *METHODS IN ENZYMOLOGY, **275***, 328-347.

Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kanegae, T., Ogura, Y., Kohchi, T., Ohyama, K. (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial-DNA - A primitive form of plant mitochondrial genome. *JOURNAL OF MOLECULAR BIOLOGY, **223***(1), 1-7

Oeser, B. and Tudzynski, P. (1989). The linear mitochondrial plasmid pCLK1 of the phytopathogenic fungus *Claviceps purpurea* may code for a DNA polymerase and an RNA polymerase. *MOLECULAR AND GENERAL GENETICS.* **217**, 132-140.

Osiewacz, H., Hermanns, J., Marcou, D., Triffi, M., Esser, K. (1989) Mitochondrial DNA rearrangements are correlated with a delayed amplification of the mobile intron (PLDNA) in a long-lived mutant of *Podospora anserina. MUTATION RESEARCH. **219***(1), 9-15

Ostrove, J., Rosenfeld, P., Williams, J., & Kelly, T. (1983). *In vitro* complementation as an assay for purification of adenovirus DNA replication proteins. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA-BIOLOGICAL SCIENCES, **80***(4), 935-939.

Pakula, T., Caldentey, J., Gutierrez, C., Olkkonen, V., Salas, M., & Bamford, D. (1993). Overproduction, purification, and characterization of DNA binding protein p19 of bacteriophage PRD1. *GENE, **126*** (1), 99-104.

Pakula, T., Caldentey, J., Serrano, M., Gutierrez, C., Hermoso, J., Salas, M., & Bamford, D. (1990). Characterization of a DNA binding protein of bacteriophage PRD1 involved in DNA replication. *NUCLEIC ACIDS RESEARCH, **18***(22), 6553-6557.

Petit, C., Rigg, G., Pazzani, C., Smith, A., Sieberth, V., Stevens, M., Boulnois, G., Jann, K., Roberts, I. (1995) Region-2 of the *Escherichia coli* k5 capsule gene cluster encoding proteins for the biosynthesis of the k5 polysaccharide. *MOLECULAR MICROBIOLOGY. **17***(4), 611-620.

Pincus, S., & Rekosh, D. (1984). The inhibition of both initiation and elongation of adenovirus DNA replication by actinomycin-D. *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, **121***(3), 1033-1041.

Pincus, S., Robertson, W., & Rekosh, D. (1981). Characterization of the effect of aphidicolin on adenovirus DNA replication - evidence in support of a protein primer model of initiation. *NUCLEIC ACIDS RESEARCH, **9***(19), 4919-4938.

Pring, D.R., III, L., Hu, W.W.L. and Timothy, D.H. (1977). Unique DNA associated with mitochondria in the "S"-type cytoplasm of male sterile maize. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE OF THE UNITED STATES OF AMERICA* **74**, 2904-2908.

Pronk, R., Stuiver, M., & Vandervliet, P. (1992). Adenovirus DNA replication - the function of the covalently bound terminal protein. *CHROMOSOMA, **102***(1 SS), S 39-S 45.

Pronk, R., Vandriel, W., & Vandervliet, P. (1994). Replication of adenovirus DNA *in vitro* is ATP-independent. *FEBS LETTERS, **337***(1), 33-38.

Pruijn, G., Vandriel, W., Vanmiltenburg, R., & Vandervliet, P. (1987). Promoter and enhancer elements containing a conserved sequence motif are recognized by nuclear factor-III, a protein stimulating adenovirus DNA replication. *EMBO JOURNAL, **6***(12), 3771-3778.

Puehler, G., Lottspeich, F., Zillig, W. (1989) Organization and nucleotide sequence of the genes encoding the large subunit- A, subunit-B and subunit-C of the DNA dependent RNA-polymerase of the archaebacterium *Sulfolobus acidocaldarius. NUCLEIC ACIDS RESEARCH. **17***(12), 4517-4534

Ramachandra, M., Sasaguri, Y., Nakano, R., & Padmanabhan, R. (1996). Heterologous expression, purification, and characterization of adenovirus DNA polymerase and preterminal protein. *METHODS IN ENZYMOLOGY, **275***, 168-194.

Rappuoli, R. & Covacci, A. (1996) CAG, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA. **93***(25), 14648-14653.

Reay, P. (1997) Purification and molecular analysis of the *Kluyveromyces lactis* linear killer plasmids. Ph.D. Thesis, University of Leicester, UK.

Rodriguez, J., Kahn, J., Esteban, M. (1986) *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF AMERICA. **83***, 9566-9570

Rohe, M., Schründer, J., Tudzynski, P., & Meinhardt, F. (1992). Phylogenetic relationships of linear, protein-primed replicating genomes. *CURRENT GENETICS, **21***, 173-176.

Rohe, M., Schrage, K. and Meinhardt, F. (1991). The linear plasmid pMC3-2 from *Morchella conica* is structurally related to adenoviruses. *CURRENT GENETICS* **20**, 527-533.

Romanos, M., & Boyd, A. (1988). A transcriptional barrier to expression of cloned toxin genes of the linear plasmid k1 of *Kluyveromyces lactis* - evidence that native k1 has novel promoters. *NUCLEIC ACIDS RESEARCH, 16*(15), 7333-7350.

Roovers, D., Vanderlee, F., Vanderwees, J., & Sussenbach, J. (1993). Analysis of the adenovirus type-5 terminal protein-precursor and DNA polymerase by linker insertion mutagenesis. *JOURNAL OF VIROLOGY, 67*(1), 265-276.

Rosenfeld, P., Wides, R., Oneill, E., Challberg, M., & Kelly, T. (1986). Adenovirus DNA replication *in vitro* - characterization of the DNA and protein-requirements. *JOURNAL OF CELLULAR BIOCHEMISTRY,* (**S10B**), 240-240.

Rosmorduc, O., Petit, M., Pol, S., Capel, F., Bortolotti, F., Berthelot, P., Brechot, C., & Kremsdorf, D. (1995). *In vivo* and *in vitro* expression of defective hepatitis B virus- particles generated by spliced hepatitis B virus-RNA. *HEPATOLOGY, 22*(1), 10-19.

Salas, M., Freire, R., Soengas, M., Esteban, J., Mendez, J., Bravo, A., Serrano, M., Blasco, M., Lazaro, J., Blanco, L., Gutierrez, C., & Hermoso, J. (1995). Protein nucleic acid interactions in bacteriophage ϕ29 DNA replication. *FEMS MICROBIOLOGY REVIEWS, 17*(1-2), 73-82.

Salas, M. (1991). Protein priming of DNA replication. *ANNUAL REVIEW OF BIOCHEMISTRY* **60**, 39-71.

Salmon, K., & Hajahmad, Y. (1994). Sequence-analysis of bovine adenovirus type-2 early region-1 and *pix* gene. *INTERVIROLOGY, 37*(5), 298-305.

Samac, D.A. and Leong, S.A. (1988). Two linear plasmids in mitochondria of *Fusarium solani* f. sp. *curcurbitae. PLASMID* **19**, 57-67.

Sato, A., Figueroa, F., Ohuigin, C., Reznick, D., Klein, J. (1996) Identification of major histocompatibility complex genes in the guppy, *Poecilia reticulata. IMMUNOGENETICS. 43*(1-2), 38-49.

Saumitoulaprade, P., Pannenbecker, G., Boutinstadler, V., Michaelis, G., & Vernet, P. (1991). Plastid DNA diversity in natural-populations of *Beta maritima* showing additional variation in sexual phenotype and mitochondrial-DNA. *THEORETICAL AND APPLIED GENETICS, 81*(4), 533-536.

Savilahti, H., & Bamford, D. (1987). The complete nucleotide-sequence of the left very early region of *Escherichia coli* bacteriophage PRD1 coding for the terminal protein and the DNA polymerase. *GENE, 57*(1), 121-130.

Savilahti, H., & Bamford, D. (1993). Protein-primed DNA replication - role of inverted terminal repeats in the *Escherichia coli* bacteriophage-PRD1 life-cycle. *JOURNAL OF VIROLOGY, 67*(8), 4696-4703.

Savilahti, H., Caldentey, J., & Bamford, D. (1989). Bacteriophage PRD1 terminal protein - expression of gene-VIII in *Escherichia coli* and purification of the functional p8 product. *GENE, **85***(1), 45-51.

Savilahti, H., Caldentey, J., Lundstrom, K., Syvaoja, J., & Bamford, D. (1991). Overexpression, purification, and characterization of *Escherichia coli* bacteriophage-PRD1 DNA polymerase - *in vitro* synthesis of full- length PRD1 DNA with purified proteins. *JOURNAL OF BIOLOGICAL CHEMISTRY, **266***(28), 18737-18744.

Schaack, J., Guo, X., Ho, W., Karlok, M., Chen, C., & Ornelles, D. (1995). Adenovirus type-5 precursor terminal protein-expressing 293 and HeLa- cell lines. *JOURNAL OF VIROLOGY, **69***(7), 4079-4085.

Schaack, J., Guo, X., & Langer, S. (1996). Characterization of a replication-incompetent adenovirus type-5 mutant deleted for the pre-terminal protein gene. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, **93***(25), 14686-14691.

Schaffrath, R., Stark, M., Gunge, N., & Meinhardt, F. (1992). *Kluyveromyces lactis* killer system - ORF1 of pGKL2 has no function in immunity expression and is dispensable for killer plasmid replication and maintenance. *CURRENT GENETICS, **21***(4-5), 357-363.

Schaffrath, R., Soond, S., & Meacock, P. (1995). The DNA and RNA polymerase genes of yeast plasmid pGKL2 are essential loci for plasmid integrity and maintenance. *MICROBIOLOGY-UK, **141***(Pt10), 2591-2599.

Schaffrath, R., & Meacock, P. (1995). *Kluyveromyces lactis* killer plasmid pGKL2 - molecular analysis of an essential gene, ORF5. *YEAST, **11***(7), 615-628.

Schaffrath, R., Soond, S., & Meacock, P. (1995). Cytoplasmic gene-expression in yeast - a plasmid-encoded transcription system in *Kluyvaromyces lactis*. *BIOCHEMICAL SOCIETY TRANSACTIONS, **23***(1), S 128-S 128.

Schaffrath, R., & Meacock, P. (1996). A cytoplasmic gene-shuffle system in *Kluyveromyces lactis* - use of epitope tagging to detect a killer plasmid-encoded gene-product. *MOLECULAR MICROBIOLOGY, **19***(3), 545-554.

Schaffrath, R., Meinhardt, F., & Meacock, P. (1996). Yeast killer plasmid pGKL2 - molecular analysis of UCS5, a cytoplasmic promoter element essential for ORF5 gene-function. *MOLECULAR & GENERAL GENETICS, **250***(3), 286-294.

Schaffrath, R., Meinhardt, F., & Meacock, P. (1997). ORF7 of yeast plasmid pGKL2: analysis of gene expression *in vivo*. *CURRENT GENETICS, **31***(2), 190-192.

Schmid, S., & Hearing, P. (1997). Bipartite structure and functional independence of adenovirus type 5 packaging elements. *JOURNAL OF VIROLOGY, 71*(5), 3375-3384.

Schrunder, J., & Meinhardt, F. (1995). An extranuclear expression system for analysis of cytoplasmic promoters of yeast linear killer plasmids. *PLASMID, 33*(2), 139-151.

Schrunder, J., Gunge, N., & Meinhardt, F. (1996). Extranuclear expression of the bacterial xylose isomerase (*xyla*) and the UDP-glucose dehydrogenase (*hasb*) genes in yeast with *Kluyveromyces lactis* linear killer plasmids as vectors. *CURRENT MICROBIOLOGY, 33*(5), 323-330.

Senkevich, T., Bugert, J., Sisler, J., Koonin, E., Darai, G., Moss, B. (1996) . *SCIENCE. 273,* 813-816.

Shiue, S., Hsieh, J., & Ito, J. (1991). Mapping of the DNA linking tyrosine residue of the PRD1 terminal protein. *NUCLEIC ACIDS RESEARCH, 19*(14), 3805-3810.

Shepherd, H., Ligon, J., Bolen, P., Kurtzman, C. (1987) Cryptic DNA plasmids of the heterothallic yeast *Saccharomycopsis crataegensis. CURRENT GENETICS. 12,* 297-304.

Sinclair, J., Stevens, B., Sanghav, I., Rabinowitz, M. (1967) *SCIENCE.* **156,** 1234-1237.

Sinha, K., Keen, J., Ogun, S., Holder, A. (1996) Comparison of two members of a multigene family coding for high-molecular mass *rhoptry* proteins of *Plasmodium yoelii. MOLECULAR BIOCHEMISTRY OF PARASITOLOGY. 76.* 329-332

Smith, D., Johnson, K. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *GENE 67*, 31-40

Snounou, G., Viriyakosol, S., Holmquist, G., Berzins, K., Perlmann, P. (1988) Cloning of genomic fragment from *Plasmodium chabaudi* expressing a 105 kilodalton antigen epitope. *MOLECULAR AND BIOCHEMICAL PARASITOLOGY, 28*(2), 153-161.

Soengas, M., Gutierrez, C., & Salas, M. (1995). Helix-destabilizing activity of ϕ29 single-stranded-DNA binding- protein - effect on the elongation rate during strand displacement DNA replication. *JOURNAL OF MOLECULAR BIOLOGY, 253*(4), 517-529.

Soond, S.M. (1994). The genetic manipulation of the linear plasmids of *Kluyveromyces lactis*. Ph.D. Thesis, University of Leicester, UK.

Sor, F., Wesolowski, M. and Fukuhara, H. (1983). Inverted terminal repititions of the two linear DNAs associated with the killer character of the yeast *Kluyveromyces lactis. Nucleic Acids Res.* **11**, 5037-5044.

Sor, F., & Fukuhara, H. (1985). Structure of a linear plasmid of the yeast *Kluyveromyces lactis* - compact organization of the killer genome. *CURRENT GENETICS, 9*(2), 147-155.

Stam, J., Kwakman, J., Meijer, M., & Stuitje, A. (1986). Efficient isolation of the linear DNA killer plasmid of *Kluyveromyces lactis* - evidence for location and expression in the cytoplasm and characterization of their terminally bound proteins. *NUCLEIC ACIDS RESEARCH, 14*(17), 6871-6884.

Stanglmaier, M., & Winnacker, E. (1996). Cloning, sequence determination and functional expression of the genes encoding adenovirus type-4 polymerase and the terminal protein- precursor. *GENE, 168*(2), 177-182.

Stark, M.J.R., Mileham, A.J., Romanos, M. and Boyd, A. (1984). Nucleotide sequence and transcriptional analysis of a linear DNA plasmid associated with the killer character of the yeast *Kluyveromyces lactis*. *NUCLEIC ACIDS RESEARCH.* **12**, 6011-6033.

Stark, M. J. R., & Boyd, A. (1986). The killer toxin of *Kluyveromyces lactis*: characterization of the toxin subunits and identification of the genes which encode them. *THE EMBO JOURNAL,* 1995-2002.

Stark, M. J., Boyd, A., Mileham, A. J., & Romanos, M. A. (1990). The Plasmid-encoded Killer System of *Kluyveromyces lactis*: A Review. *YEAST, 6*, 1-29.

Stillman, B. (1981). Adenovirus DNA replication *in vitro* - a protein linked to the 5' end of nascent DNA strands. *JOURNAL OF VIROLOGY, 37*(1), 139-147.

Stillman, B., Topp, W., & Engler, J. (1982). Conserved sequences at the origin of adenovirus DNA replication. *JOURNAL OF VIROLOGY, 44*(2), 530-537.

Sussenbach, J., & Vandervliet, P. (1983). The mechanism of adenovirus DNA replication and the characterization of replication proteins. *CURRENT TOPICS IN MICROBIOLOGY AND IMMUNOLOGY, 109*, 53-73.

Szewczyk, B., Summers, D. (1992) Efficient elution of purified proteins from Polyvinylidene Diflouride membranes (Imobilon) after transfer from SDS-PAGE and their use as immunogens. In: Methods in Molecular Biology, Volume 10.

Szostak, J., & Blackburn, E. (1982). Cloning yeast telomeres on linear plasmid vectors. *CELL, 29*(1), 245-255.

Takano, H., Kawano, S., Kuroiwa, T. (1994) Complex terminal structure of a linear mitochondrial plasmid from *Physarum polycephalum* - 3' terminal inverted repeats and an ORF encoding DNA polymerase. *CURRENT GENETICS. 25*(3) 252-257.

Takeda, M., Hiraishi, H., Takesako, T., Tanase, S., & Gunge, N. (1996). The Terminal Protein of the Linear DNA Plasmid pGKL2 Shares an N-terminal Domain of the Plasmid-Encoded DNA Polymerase. *YEAST, **12***, 241-246.

Tamanoi, F., & Stillman, B. (1983). Initiation of adenovirus DNA replication *in vitro* requires a specific DNA sequence. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA-BIOLOGICAL SCIENCES, **80***(21), 6446-6450.

Tamanoi, F., & Stillman, B. (1983). The origin of adenovirus DNA replication. *CURRENT TOPICS IN MICROBIOLOGY AND IMMUNOLOGY, **109***, 75-87.

Tanguy-Rougeau, C., Chen, X.J., Wesolowski-Louvel, M. and Fukuhara, H. (1990). Expression of a foreign Km^R gene in linear killer DNA plasmids in yeast. *GENE* **91**, 43-50.

Tartaglia, J., Winslow, J., Goebel, S., Johnson, G., Taylor, J., Paoletti, E. (1990) Nucleotide sequence analysis of a 10.5 kbp *Hind*III fragment of fowlpox virus - relatedness to the central portion of the *Vaccinia* virus *Hind*III D-region. *JOURNAL OF GENERAL VIROLOGY. **70***(7) 1517-1524.

Thiele, D., Willems, H., Haas, M., Krauss, H. (1995) Analysis of the entire nucleotide sequence of the cryptic plasmid QPH1 from *Coxiella burnetti. EUROPEAN JOURNAL OF EPIDEMIOLOGY. **10***, 413-420.

Thompson, A., & Oliver, S. (1986). Physical separation and functional interaction of *Kluyveromyces lactis* and *Saccharomyces cerevisiae* ARS elements derived from killer plasmid DNA. *YEAST, **2***(3), 179-191.

Tokunaga, T., Wada, N. and Hishinuma, F. (1987a). Expression and identification of immunity determinants on the linear DNA killer plasmids pGKL1 and pGKL2 in *Kluyveromyces lactis. NUCLEIC ACIDS RESEARCH.* **15**, 1031-1046.

Tokunaga, T., Wada, N. and Hishinuma, F. (1987b). A novel yeast secretion vector utilizing signal of killer toxin encoded on the yeast linear plasmid pGKL1. *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNITY.* **114**, 613-619.

Tokunaga, T., Wada, N. and Hishinuma, F. (1988). A novel yeast secretion signal isolated from 28K killer precursor encoded on the linear plasmid pGKL1. *NUCLEIC ACIDS RESEARCH.* **16**, 7499-7511.

Tokunaga, T., Kawawura, A. and Hishinuma, F. (1989). Expression of pGKL1 killer 28K subunit in *Saccharomyces cerevisiae*: identification of 28K subunit as a killer protein. *NUCLEIC ACIDS RESEARCH.* **17**, 3435-3446.

Tokunaga, T., Kawawura, A., Kitada, K. and Hishinuma, F. (1990). Secretion of killer toxin encoded on the linear DNA plasmid pGKL1 from *Saccharomyces cerevisiae*. *JOURNAL OF BIOLOGICAL CHEMISTRY*. **265**, 17274-17280.

Tokunaga, T., Kawawura, A., Omori, A. and Hishinuma, F. (1991). Purification and determination of the NH_2-terminal amino acid sequence of mouse α-amylase secreted from *Saccharomyces cerevisiae*: correct processing of the secretion signal from pGKL1 killer 28kDa precursor protein. *BIOCHIMICA BIOPHYSICA ACTA* **1080**, 135-137.

Tokunaga, T., Kawawura, A., Omori, A. and Hishinuma, F. (1992a). Structure of yeast pGKL1 128 kDa killer toxin secretion signal sequence: processing of the 128 kDa killer-toxin secretion signal α-amylase fusion protein. *EUROPEAN JOURNAL OF BIOCHEMISTRY* **203**, 415-423.

Tokunaga, T., Kawawura, A., Omori, A. and Hishinuma, F. (1992b). Characterisation of mouse α–amylase secreted from *Saccharomyces cerevisiae* by the pGKL1 128kDa killer secretion signal. *BIOSCI. BIOTECH. BIOCHEM.* **56**, 155-156.

Tokunaga, T., Kawawura, A., Yonekyu, S., Kishida, M. and Hishinuma, F. (1993). Secretion of mouse α-amylase from fission yeast *Schizosaccharomyces pombe*: Presence of chymostatin-sensitive protease activity in the culture medium. *YEAST* **9**, 379-387.

Tomb, J., White,O. Kerlavage,A., Clayton,R., Sutton,G., Fleischmann,R., Ketchum,K., Klenk,H., Gill,S., Dougherty,B., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H., Glodek, A., McKenney,K., Fitzegerald,L., Lee,N., Adams,M., Hickey,E., Berg,D., Gocayne, J., Utterback, T., Peterson, J., Kelley, J., Cotton, M., Weldman, J., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W., Weidman, J., Fujii, C., Borodovsky, M., Karp, P., Smith, H., Fraser, C., Venter, J. (1997) . The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *NATURE*. ***388***(6642), 539-547.

Tommasino, M., Ricci, S., & Galeotti, C. (1988). Genome organization of the killer plasmid pGKL2 from *Kluyveromyces lactis*. *NUCLEIC ACIDS RESEARCH*, ***16***(13), 5863-5878.

Tommasino, M. (1991). Killer System of *Kluyveromyces lactis*: the Open Reading Frame 10 of the pGKL2 Plasmid Encodes a Putative DNA Binding Protein. *YEAST*, *7*, 245-252.

Tuma, R., Bamford, D., & Thomas, G. (1994). Structure, interactions and thermostability of the membrane-packaged dsDNA genome of bacteriophage PRD1. *BIOPHYSICAL JOURNAL*, ***66***(2 Pt2), A 296-A 296.

Tuma, R., Bamford, J., Bamford, D., & Thomas, G. (1996). Structure, interactions and dynamics of PRD1 virus 2: Organization of the viral membrane and DNA. *JOURNAL OF MOLECULAR BIOLOGY, 257*(1), 102-115.

Vanleeuwen, H., Strating, M., Cox, M., Kaptein, R., & Vandervliet, P. (1995). Mutation of the OCT-1 POU-specific recognition helix leads to altered DNA binding and influences enhancement of adenovirus DNA replication. *NUCLEIC ACIDS RESEARCH, 23*(16), 3189-3197.

Vanleeuwen, H., Rensen, M., & Vandervliet, P. (1997). The OCT-1 POU homeodomain stabilizes the adenovirus preinitiation complex via a direct interaction with the priming protein and is displaced when the replication fork passes. *JOURNAL OF BIOLOGICAL CHEMISTRY, 272*(6), 3398-3405.

Volkert, F. C., Wilson, D. W., & Broach, J. R. (1989). Deoxyribonucleic Acid Plasmids in Yeasts. *Microbiological Reviews, 53*(3), 299-317.

Vrati, S., Brookes, D., Boyle, D., & Both, G. (1996). Nucleotide-sequence of ovine adenovirus tripartite leader sequence and homologs of the *iva2*, DNA polymerase and terminal proteins. *GENE, 177*(1-2), 35-41.

Vydelingum, S., Baylis, S., Bristow, C., Smith, G., Dixon, L. (1993) Duplicated genes within the variable right end of the genome of a pathogenic isolate of African swine fever virus. *JOURNAL OF GENERAL VIROLOGY. 74*(10), 2125-2130

Watabe, K., Ito, J. (1983) A novel DNA-polymerase induced by *Bacillus subtilis* phage PHI-29. *NUCLEIC ACIDS RESEARCH. 11*(23) 8333-8342

Weber, J., & Tihanyi, K. (1994). Adenovirus endopeptidases. *METHODS IN ENZYMOLOGY, 244*, 595-604.

Weber, M., Bronsema, V., Bartos, H., Bosserhoff, A., Bartenschlager, R., & Schaller, H. (1994). Hepadnavirus p-protein utilizes a tyrosine residue in the TP domain to prime reverse transcription. *JOURNAL OF VIROLOGY, 68*(5), 2994-2999.

Webster, A., Leith, I., & Hay, R. (1994). Activation of adenovirus-coded protease and processing of preterminal protein. *JOURNAL OF VIROLOGY, 68*(11), 7292-7300.

Webster, A., Leith, I., & Hay, R. (1997). Domain organization of the adenovirus preterminal protein. *JOURNAL OF VIROLOGY, 71*(1), 539-547.

Wesolowski, M., Algeri, A., Goffrini, P., Fukuhara, H. (1982) Killer DNA plasmids of the yeast *Kluyveromyces lactis*. I. Mutations affecting killer phenotype. *CURRENT GENETICS. 5*, 191-197

Wesolowski, M., Algeri, A., Fukuhara, H. (1982) Killer DNA plasmids of the yeast *Kluyveromyces lactis*. III. Plasmid recombination. *CURRENT GENETICS*. **5**, 205-208.

Wesolowski-Louvel, M., & Fukuhara, H. (1990). A palindromic mutation of the linear killer plasmid-k2 of yeast. *NUCLEIC ACIDS RESEARCH, 18*(16), 4877-4882.

Wickner, R. (1981). [Hok], a new plasmid that helps a replication-defective k1-killer plasmid mutant. *PLASMID, 5*(2), 230-231.

Wickner, R., & Tohe, A. (1982). *Hok*, a new yeast non-mendelian trait, enables a replication- defective killer plasmid to be maintained. *GENETICS, **100***(2), 159-174.

Wilber, K., Proffitt, J. (1987) UCLA Yeast Symposium, USA, 106

Wilson, D.W. and Meacock, P.A. (1988). Extranuclear gene expression in yeast: evidence for a plasmid-encoded RNA polymerase of unique structure. *NUCLEIC ACIDS RESEARCH*. **16**, 8097-8110.

Wilson, J.H., Leung, W.Y., Bosco, G., De Dieu and Haber, J.E. (1994). The frequency of gene targeting in yeast depends on the number of target copies. *GENETICS* **91**, 177-181.

Yanez, R., Rodriguez, J., Nogal, M., Yuste, L., Enriquez, C. Rodriguez, J., Vinuela, E. (1995) Analysis of the complete nucleotide sequence of African swine fever virus. *VIROLOGY. **208***(1). 249-278

Yoo, S., & Ito, J. (1989). Protein-primed replication of bacteriophage-PRD1 genome *in vitro*. *VIROLOGY, **170***(2), 442-449.

Yoo, S., & Ito, J. (1991). Initiation of bacteriophage-PRD1 DNA replication on single-stranded templates. *JOURNAL OF MOLECULAR BIOLOGY, **222***(2), 127-131.

Yoo, S., & Ito, J. (1991). Sequence requirements for protein-primed DNA replication of bacteriophage PRD1. *JOURNAL OF MOLECULAR BIOLOGY, **218***(4), 779-789.

Yui, Y., Katayose, Y., & Shishido, K. (1988). 2 linear plasmid-like DNA elements simultaneously maintained in *Pleurotus ostreatus*. *BIOCHIMICA ET BIOPHYSICA ACTA, **951***(1), 53-60.

Zhu, W., & Ito, J. (1994). Family-A and family-B DNA polymerases are structurally related - evolutionary implications. *NUCLEIC ACIDS RESEARCH, **22***(24), 5177-5183.

Zhu, W., & Ito, J. (1994). Purification and characterization of PRD1 DNA polymerase. *BIOCHIMICA ET BIOPHYSICA ACTA-GENE STRUCTURE AND EXPRESSION, **1219***(2), 267-276.

Zhu, W., Leavitt, M., Jung, G., & Ito, J. (1994). Mutagenesis of a highly conserved lysine-340 of the PRD1 DNA-polymerase. *BIOCHIMICA ET BIOPHYSICA ACTA-GENE STRUCTURE AND EXPRESSION, **1219***(2), 260-266.

Zoulim, F., & Seeger, C. (1994). Reverse transcription in hepatitis B viruses is primed by a tyrosine residue of the polymerase. *JOURNAL OF VIROLOGY, **68***(1), 6-13.