Genome organization of the killer plasmid pGK12 from *Kluyveromyces lactis*

M.Tommasino, S.Ricci and C.L.Galeotti

Slavo Research Centre, Via Fiorentina 1, 53100 Siena, Italy

Received March 24, 1988; Revised and Accepted June 9, 1988

Accession no.X07776

ABSTRACT

We have determined the entire sequence of the plasmid K2 from *Kluyveromyces lactis* which is involved in the maintenance of both killer plasmids in the cell. K2 shares many of the characteristics of the smaller killer plasmid K1: high A+T content (74.7%) and very compact genomic organization. K2 contains ten open reading frames. Some of them overlap on different strands and some on the same strand. Northern blotting of K2 transcripts shows that at least eight ORFs are transcribed. Analysis of the predicted aminoacid sequence of ORF 2 from K2 reveals homology with the aminoacid sequence of ORF 1 from K1 and with several viral DNA polymerases. The sequence of K2 from *Saccharomyces cerevisiae* F102-2 was also determined. Only one nucleotide difference was found between the K2 sequence from the two yeasts. This mutation does not change the genome organization of the plasmid and has only minimal effect on the structure of the encoded proteins.

INTRODUCTION

The presence of the two linear plasmids, termed pGK11 (8.9 Kb) and pGK12(13.4 Kb) (or K1 and K2, respectively, for short), in strains of *Kluyveromyces lactis* confers the killer phenotype to the host cells, by promoting the secretion of a toxin able to inhibit the growth of sensitive strains (1,2). Both plasmids can be transferred from a killer strain of *K. lactis* to *Saccharomyces cerevisiae* by protoplast fusion or transformation (3,4). The plasmids are stably maintained and expressed in *S. cerevisiae*. Both plasmids have terminal inverted repeats (5) and their 5' ends are blocked by a protein with apparent M.W.s of 36 Kd (K2) and 28 Kd (K1) (6). The function of these terminal proteins is still unknown, although similar systems have been described in which terminally attached proteins function as primers in DNA replication, e.g. adeno-virus (for review, see 7).

Three groups have independently sequenced the smaller plasmid K1, two of them using the original K. lactis strain (8,9) and the other using an S. cerevisiae transformant (F102-2) (10). Four and five open reading frames (ORFs) were found in K. lactis and S. cerevisiae respectively. ORFs 2 and 4 from the K. lactis K1 plasmid code for the three subunits of the killer toxin (11), while ORF 3 codes for a protein involved in the immunity system (12). Nothing is known about the function of the protein coded for by ORF 1, although some homology has been found with viral DNA polymerases (13).

Until now no mutants have been found containing only K1. However, several mutants have been isolated harboring only K2. The failure of K1 to exist alone suggests that the proteins encoded by K2 have a crucial role in the maintenance of both plasmids.

In order to understand the mechanism of replication and the gene organization of the K2 plasmid, we have determined the nucleotide sequence of K2 from K. lactis CBS 2359 and also from S. cerevisiae F102-2. Apart from one minor difference, the sequence of K2 from both organisms is identical. As in K1, the K2 genome has a high A+T content (74.7%) and a very compact structure. Analysis of the sequence shows the existence of 10 potential ORFs. Preliminary examination of the transcripts by Northern blot analysis suggests that at least eight of these are transcribed.

MATERIALS AND METHODS

Strain

K. lactis CBS 2359 was obtained from the Centraal Bureau voor Schimmelcultures and S. cerevisiae F 102-2 (Mat a, leu 2-3, -112, his 4-519, can 1 [K2+ K1+]) was kindly provided by Dr. Norio Gunge.

K2 DNA preparation

K2 DNA was purified from cultures of K. lactis CBS 2359 or S. cerevisiae F102-2 according to Stam et al. (6).

Nucleotide sequence determination

On the basis of previously published restriction maps (14), fragments of the K2 plasmid obtained after digestion with

restriction enzymes (Eco RI, Eco RI-Xho I, Bgl II, Xba I, Xba I-Xho I) were cloned into the vectors pEMBL 9, pEMBL 18 and Bluescript M13+. All Eco RI fragments from S. cerevisiae F102-2 were sequenced at least once on both strands by dideoxy chain terminator sequencing (15), using as primers synthetic 20-mer oligonucleotides. The sequence at the junctions between Eco RI fragments was determined by sequencing clones of Bgl II, Xba I and Xba I-Xho I restriction fragments. The sequence of K2 from K. lactis CBS 2359 was performed using the same cloning strategy as described above. The Eco RI fragments were sequenced on one strand. The difference between the k2 sequence from the two yeast species was confirmed by sequencing the same region on both strands from three independently isolated Eco RI clones.

Computer analysis was carried out using the UWGCG sequence analysis software package (16).

Preparation and Northern blot analysis of RNA

Yeast cells were grown to mid log-phase at 30°C in a medium containing 1% yeast extract, 2% peptone and 2% glucose. RNA was prepared by disruption of the cells with glass beads in 0.32M sucrose, 20mM Tris/HCl pH 7.5, 10mM EDTA and 0.5 mg/ml Heparin at 0°C. After the addition of 1% of sodium dodecyl sulphate, the suspension was extracted once with phenol for 5 min at 60°C and once with phenol/chloroform (1:1). The RNA in the aqueous phase was precipitated in ethanol in the presence of 0.1M NaCl (17). For Northern blot analysis, total RNA was denaturated in 50% deionized formamide, 5.5% formaldehyde in MOPS buffer (20mM MOPS, 5mM Sodium Acetate, 1mM EDTA, pH 7.0 adjusted with acetic acid) and separated in 1.5% agarose gels in the presence of 5.5% formaldehyde (18). The RNA was transferred to Gene Screen membrane (Dupont) in 25mM phosphate buffer pH 6.5. Prehybridization and hybridization of the filter were carried out in 50% deionized formamide and in 10% dextran sulphate at 42°C as described in the Gene Screen handbook. Double stranded DNA probes were labelled using random primers to a specific activity of 1-2 10^9 cpm/ μ g DNA (19).

Transcript lengths were estimated using HeLa rRNA, yeast rRNA, HeLa 7SK (20) and E. coli tRNA as size markers.

5866

[illegible]

```

12361 ATATGTAAATGGATTACCAATAGATGAAGAGAAATATTCATATTCTTTATAAAGAAAAATATAAAATATCAATGAAGATGAAAAATTTATAGAGTGAATTATATA
  Y V N G L P I D E E E N I P Y S F I K E K Y K N I K Y H E D E K I I E V D Y M N
12481 CAATGTAGTAAACAAAAGATATGCCAGATATTACTGTTATATATACAAATTAATACAAAAGATAGAGAACCTGTTAAATTACCAAAAATAGAGCAAAATTTATATAAAGAAA
  N V V K Q K D M A E Y Y C Y N N K L I Q K I E E P V K L P K I E A Q I L N K R N
12601 TGAAGAGTATAGAGTATTATATAGATAACGAATGGAAAACTTAGCTGGTAAATAGATAGATGGAAATTCCTACTCTTCCATGGAACATAAACATATATATACATGACTATAAATTAGG
  E R V E V L I D N E W K N L A G N R I D G I P T L P M N Y K H I I H D Y N K L G
12721 TAGATTAAAGAGGAGAAAATATGGATTTAACACATATAGGAAGATGAATGAAAATGGAACTGATATTATTACTTGGCCATATATCTGATGACCTATACATAGGAAGATCTATAGAAAC
  R L K K G K I M D L T H I G I V N E N G T D I I T M P Y T D D L Y I G R S I E T
12841 AAAAGATTATATACCTTTCAATATACAGAAAATGTAGAATATATGATAGCTCATGACATGTAATGGCTAATAAACAGCCAGAAAACTCATAAACCCGATCAAAAAGATTATT
  K R L I T F K Y Q K N V E I Y D K L I D M *
                                     —ORF10—▶ H A N K Q A E K L I T A I K K D Y L
12961 TAAAGAAATATATAAGAAATTTGAAGAAATAGATATAGATAGAGAAGATTTATTTGAAAATATAAGGAAGAAAAACCAAGAAAAAGAGAAATCTCCAAAAATTCCTATGAATA
  K E I I K K I E E L D I D K K D V I V E K L K E E K R N A P K I P L N K
13081 AACAATGACTATAAGAAACAGCTAGTAAAGTAAATGTACCGTAGCAGCTTGTACACCATATATGCTGGGCACATATGAACAAAACACAGAGAAATGATATAGATTACTCAAAATCAG
  Q C T K E T A S K G K C T V A A C Y N H I C M A H M N K T Q R N E Y R L L K S V
13201 TAGACATTAAACAAATTTAAATTTATATATACATTTAACAATGATGCTACTATATAGTATTTTGGGATTATGGAATATATAGAAAAGATATGACCTAATGACTCCG
  D I K T I A
                                     XhoI
                                     ↳ IR
13321 GCGAAAAGGGCGTACTCTGACATATGACCTAACATATAAAAACACGTAACACATCTTTGGGTGCTTTTGTGACCTGATTTCACATTTTATTAGCTTAAAAACCCAACTTTATTTT
  13441 TATATATATATACCTTTT 13457

```

Figure 1: Nucleotide sequence of the K2 genome.

The sequence of K2 (13457 pb) is shown together with the predicted aminoacid sequences of the ten ORFs. The ends of the inverted repeats are indicated by the arrows. The sequence underlined (position 1-116, 13342-13457) has been previously sequenced by Sor et al. (5).

RESULTS

Structure of the K2 genome

The termini of the K2 plasmid have been sequenced previously by two groups. Sor et al. (5) have sequenced approximately 220 bp at both ends of the plasmid from *K. lactis* CBS 2359. Hishimuna et al. (10) have sequenced to the first Xba I site on the right end (1317 bp) and to the first Eco RI site on the left end (797 bp) of the plasmid from *S. cerevisiae* F102-2. The plasmid was found to contain inverted terminal repeats of 182 bp with Xho I sites 120 bp from each end.

We have determined the entire sequence of K2 from the *K. lactis* strain CBS 2359 between the two terminal Xho I sites. Defined restriction fragments were subcloned using Eco RI or Eco RI-Xho I digests and each of the clones was sequenced using synthetic oligonucleotide primers. Appropriate subclones, obtained using Bgl II, Xba I or Xba I-Xho I restriction fragments which span each of the Eco RI sites, were used to confirm the sequence around the Eco RI junctions.

The entire sequence of k2 between the two Xho I sites, including the terminal 116 bp already published (5), is presented in fig. 1. Our results are in complete agreement with the data of Sor et al. (5), while we find differences at

Figure 2: Restriction map of K2 linear plasmid and schematic representation of the organization of the K2 genome. Detailed restriction map of K2 as derived from the sequence. The restriction enzymes used for the subcloning are included. The ten ORFs are represented by the arrows, indicating the location (top or lower strand) and the polarity (5'→3') of the ORFs. The solid areas correspond to the inverted repeats.

positions 360 and 855 with the sequence published by Hishimuna et al. (10). In analogy to the K1 plasmid, the base composition of K2 is very AT-rich (38.2% A, 36.5% T, 13.2% G and 12.0% C).

Fig. 2 shows a detailed restriction map of K2 deduced from the sequence. Differences from the previously published restriction maps of K2 were found with regard to the orientation of the 3.4 kb Eco RI fragment and to the number of Eco RI and Xba I sites. The additional restriction sites were found at positions 5934, 9486 (Xba I) and 11458 (Eco RI) (14). Inspection of the K2 sequence reveals ten ORFs (Fig. 2, Table I) covering about 97% of the genome. ORF 1, 4, 5, 6, 9 and 10 are coded for by the top strand. ORFs 2, 3, 7 and 8 are coded for by the lower strand. All other potential coding regions are less than 100 amino acids long and overlap completely with ORFs 1-10. The 10 ORFs are very tightly packed. As reported for K1, K2 also has two ORFs (ORFs 8,9) on opposite strands overlapping for 15 codons at their 5' ends (fig.2). In another case, two ORFs on different strands share two bases of the start or stop codons (fig. 2 ORFs 1,2). ORFs 2-3, ORFs 4-5 and ORFs 5-6 are in different frames overlapping on the same strand. Some overlap for a few codons (ORFs 4-5, 5-6) while ORF 2 and ORF 3 overlap only for two bp (fig. 2,3).

Since the *K. lactis* K1 sequence was found to differ from

TABLE I: The ORFs of K2 and main characteristics of the predicted proteins.

	NUCLEOTIDE No.	NUMBER OF AMINOACID	MOLECULAR WEIGHT	POSSIBLE GLYCOSYLATION SITES
ORF 1	230-904	224	26741	2
ORF 2	901-3885	994	117555	6
ORF 3	3882-5666	594	70639	3
ORF 4	5702-7495	597	69161	9
ORF 5	7485-7961	158	18092	0
ORF 6	7958-10906	982	113946	7
ORF 7	10901-11299	132	15498	1
ORF 8	11302-11589	95	10973	4
ORF 9	11546-12907	453	55035	3
ORF 10	12909-13220	103	12060	1

the sequence of the same plasmid from *S. cerevisiae* F102-2 (8-10), we have also sequenced the plasmid K2 obtained from *S. cerevisiae* F102-2 using the same cloning strategy as described above. We have found only one nucleotide difference between the two sequences. In the run of 6 As at position 3885 of the *K. lactis* K2 genome there is a deletion of one A in *S. cerevisiae*.

Figure 3: The region of sequence divergence of K2 from *K. lactis* and *S. cerevisiae* F 102-2.

Part of the sequence of K2 from each strain, corresponding to the junction of ORF 2 and ORF 3 (3872-3896), is shown together with the predicted amino acid sequences. The codons of ORF 2 are boxed. The arrows indicate the polarity of the ORFs. The deleted A in the *S. cerevisiae* sequence is shown in boldface.

Figure 4: Northern blot analysis of total K2 transcripts. Total RNA from three different yeast strains was analyzed by Northern blotting. Lane 1: 10 μ g RNA from *K. lactis* (non-killer strain); lane 2: 10 μ g RNA from *S. cerevisiae* F102-2 (killer strain); lane 3: 10 μ g RNA from *K. lactis* CBS 2359 (killer strain). Cloned segments covering the entire K2 genome were 32 P labelled using random primers as described (19). The position of 26S and 18S rRNA of yeast are indicated by the arrows. The lengths of each transcript is indicated on the left; these were estimated using HeLa rRNA, yeast rRNA, HeLa 7SK and *E. coli* tRNA as size markers.

This is at the junction between ORF 2 and ORF 3. The consequence of this change is, however, minimal. As shown in Fig. 3, ORF 2 does not change at all, while ORF 3 undergoes a K $\rightarrow$ N substitution and gains an extra aminoacid (E) at the C terminus.

Analysis of K2 transcripts

We have examined the transcripts of K2 by Northern blot analysis using cloned segments covering the entire K2 genome as radioactive probes. Fig. 4 shows the results obtained with total RNA purified from the non-killer strain of *K. lactis* WM 37 (lane 1) and from the killer strains *S. cerevisiae* F102-2 (lane 2) and *K. lactis* CBS 2359 (lane 3). In lane 2 and 3 eight bands are visible in the range of about 3.2-0.4 Kb, which are consistent with the predicted sizes of the ORFs of K2. In a longer exposure two extra bands are visible between the 0.8 kb band and the 0.4 kb band. None of these bands were found in the

Figure 5: Northern blot analysis of mRNA transcribed from ORF 4, ORF 5 and ORF 6.

10 μ g of total RNA from *K. lactis* CBS 2359 was analyzed by Northern blotting using the 3.4 kb Eco RI fragment, which includes part of ORF 4, the whole ORF 5 and part ORF 6, as radioactive probe. The position of 26S and 18S rRNA of yeast are indicated by the arrows. The lengths of the transcripts were estimated using the same markers as in Fig. 5.

non-killer strain (lane 1). Comparison of lane 2 and 3 shows that there is no detectable difference in the pattern of transcripts of K2 from *K. lactis* CBS 2359 and *S. cerevisiae* F 102-2.

In order to determine whether the ORFs overlapping on the same strand were transcribed as a polycistronic RNA, we hybridized the total RNA from *K. lactis* CBS 2359 with a fragment of K2 comprising ORFs overlapping on the same strand. Fig. 5 shows the results obtained using as a probe the 3.4 kb Eco RI fragment (see Fig. 1) which covers part of ORF 4, ORF 5 and part of ORF 6. Three transcripts, corresponding to the estimated sizes of ORFs 4, 5 and 6, were detected, suggesting that each ORF is transcribed independently.

DISCUSSION

The overall sequence organization of K2 is similar to that of the smaller plasmid K1, which has been sequenced previously (8-10). The base composition in both cases is rich in A+T (73.2% and 74.7% for K1 and K2 respectively) and both plasmids have an extremely compact genome structure. In K1 four ORFs cover 95% of the entire sequence. ORF 2 and ORF 3 are on

different strands and overlap for ten codons at their 3' ends. The structure of K2 is even more compact than that of K1. Approximately 97% of the K2 sequence potentially codes for proteins. Apart from the terminal inverted repeats, only 140 bp are non-coding. There are 35 bp between ORF3 and ORF4, and the remaining ones are approximately equally distributed at each end adjacent to the inverted repeats. The codon usage of K1 and K2 is also very similar. The high percentage of A+T in K2 predicts a bias towards AT-rich codons in the ORFs of K2. In fact Ile, Leu, Lys, Asn and Glu are abundant in all ORFs of K2.

A novel feature of the organization of the ORFs in K2, compared to K1, is the presence of overlapping ORFs in different frames of the same strand (2-3, 4-5, 5-6). In all cases, overlapping ORFs on the same strand have a methionine at the beginning of the coding region. It thus seems unlikely that any of these polypeptides are expressed via a translational frameshift as found, for example, in the case of Ty elements (21,22). This conclusion is also supported by analysis of the transcripts following Northern blotting.

The results of Northern blotting (Fig. 4) show that at least eight transcripts can be detected in total RNA extracted from cells containing the killer plasmids. ORF3 and ORF4 have approximately the same size, so it is possible that their transcripts may comigrate (third band, fig. 4, lane 2-3). More detailed mapping of the transcripts using subgenomic fragments showed three independent transcripts of 3.0, 1.7 and 0.8 Kb (fig. 6) which correspond approximately to the expected mRNA sizes of ORFs 6, 4 and 5, respectively. This suggests that the translation of each of these mRNA is initiated independently even though the ORFs are overlapping on the same strand.

Comparison of the nucleotide sequence of K1 isolated from K. lactis and from S. cerevisiae (F102-2) revealed changes in the structural organization of the K1 genome. The original K1 sequence from K. lactis is 8874 bp and has four large ORFs. The sequence of K1 determined from S. cerevisiae was 2 bp longer and, as a consequence, ORF 1 was split into two separate ORFs. To assess the extent of sequence divergence of K2 between these two hosts, we have also determined the nucleotide sequence of K2 isolated from S. cerevisiae. The deletion of a single

	I	II	III
K ₁ ORF1 (648)	<u>LCLDVKS</u> <u>SLYPASMAFYDQ</u> <u>PYG</u> (783)	<u>VIKIIMN-SLWGKFA</u> (862)	<u>CIYSDTDSIFV</u>
K ₂ ORF2 (633)	<u>VYADVVS</u> <u>SLYPSAMKLLKHSYG</u> (768)	<u>VAKIALNGGGYGK</u> <u>FV</u> (874)	<u>IIYSDTDSIFV</u>
S ₁ URF3 (482)	<u>YYDVNS</u> <u>SLYPSSMLDDM-PIG</u> (615)	<u>IYKIIMN-SLYGRFG</u> (713)	<u>CYYTDTDSVVV</u>
Phi 29 (245)	<u>MVFDVNS</u> <u>LYPAQMYSRLLPYG</u> (380)	<u>LAKLMLN-SLYGKFA</u> (450)	<u>IIYCDTDSIHL</u>
EBV (583)	<u>LVVDFAS</u> <u>LYPSIIQAHNLCYS</u> (681)	<u>AIKCTCN-AVYG-FT</u> (753)	<u>VIIYGDTSLSFI</u>
HSV (714)	<u>VVFDFA</u> <u>SLYPSIIQAHNLCFS</u> (809)	<u>AIKVVCN-SVYG-FT</u> (882)	<u>IIYGDTSIFV</u>
VV (521)	<u>LIFDYN</u> <u>SLYPSNVCIFGNLSP</u> (634)	<u>TIKIVAN-SVYGLMG</u> (722)	<u>SVYGDTSVFT</u>
AD2 (536)	<u>YVYDIC</u> <u>GMYASA-LTHMPFWG</u> (693)	<u>IAKLLSN-ALYGSFA</u> (866)	<u>SVYGDTSLSFV</u>

Figure 6: Amino acid sequence homology between ORF 2 of K2 and DNA polymerases encoded by other naturally occurring episomal elements.

The predicted aminoacid sequence of ORF 2 was compared with the aminoacid sequence of: ORF 1 from K1 (8,9); ORF 3 from the maize mitochondrial linear plasmid (S1) (28) DNA polymerase of bacteriophage phi 29 (29), Epstein-Barr virus (23), Herpes simplex virus (30), Vaccinia virus (31) and Adenovirus 2 (23). Numbers in parentheses indicate the position of the three blocks of aminoacids in the protein sequence. The homologies with ORF 2 are underlined.

adenosine in S. cerevisiae (at position 3885) is the only difference between the two sequences, with minimal effect on the structure of the K2 genome (see fig. 3).

A search for homology between the proteins encoded by K1 and K2 revealed that ORF 1 of K1 and ORF 2 of K2 have 46% of conserved aminoacids. It has been shown recently that the protein encoded by ORF 1 of K1 has a significant homology with three conserved aminoacid blocks characteristic of a number of viral DNA polymerases (13,23) and with a protein encoded by a linear plasmid present in maize mitochondria (24) (fig. 6). The homology between ORF 1 (K1) and ORF 2 (K2) is particularly high in the conserved regions of the viral DNA polymerases (fig. 6), suggesting that ORF 1 of K1 and ORF 2 of K2 may encode plasmid specific DNA polymerases.

The organization of these linear plasmids of K. lactis seems similar in many respects to more extensively studied linear plasmid systems, such as Adenovirus and Bacteriophage phi 29 (for reviews, see 7,25). In the latter two systems, DNA replication occurs by a mechanism of strand displacement primed

from a nucleotide bound at the terminal protein. In Adenovirus, the DNA polymerase has been found associated with the premature form of the terminal protein (26,27). If the replication of K1 and K2 is similar to that described for Adenovirus, it is likely that each of the two plasmids require specific DNA polymerases, since K1 and K2 have different terminal proteins (6) and different terminal repeats (5).

By analogy with other linear plasmids, it is also likely that their terminal proteins are plasmid encoded products. Functions have been assigned to each of the four ORFs of K1, so it is possible that the terminal proteins for both plasmids are encoded by K2. This is also consistent with the observation that K2 is required for the maintenance of K1. The size of the terminal proteins of the two killer plasmids has been estimated to be 36 and 28 Kd for K2 and K1 respectively (6). Excluding ORF 2 (see above), ORFs 3, 4, 6, or 9 may code for a protein with the same or larger size than that of the mature form of the terminal proteins.

No speculation can be advanced on the possible function of the proteins encoded by the smaller ORFs (1, 7, 8, 10), although the very high percentage of lysine in ORF 10 (22%) suggests that its translational product may be a DNA binding protein.

The similarity between the genome structure of the killer system of K. lactis and the genome structure of a variety of animal viruses (linear DNA, terminal proteins covalently linked to the DNA, terminal inverted repeats), and the homology of the proteins encoded by K1 ORF 1 and K2 ORF 2 with viral DNA polymerases, suggests a viral origin for the K1 and K2 plasmids.

Acknowledgements

The authors are deeply indebted to Dr. Kinsey Maundrell for his constant encouragement, and helpful and stimulating discussions during the course of this work. They also thank him, together with Drs. M. Melli and R. Rappuoli, for valuable advice and critical reading of the manuscript, G. Corsi for photographic work, and S. Weemys for secretarial assistance.

REFERENCES

1. Gunge, N., Tamaru, A., Ozawa, F. and Sakaguchi, K. (1981) *J. Bacteriol.* 145, 382-390.
2. Wesolowski, M., Algeri, A., Goffrini, P. and Fukuhara, H. (1982) *Current Genet.* 5, 191-197.
3. Gunge, N. and Sakaguchi, K. (1981) *J. Bacteriol.* 147, 155-160.
4. Gunge, N., Murata, K. and Sakaguchi, K. (1982) *J. Bacteriol.* 151, 462-464.
5. Sor, F., Wesolowski, M. and Fukuhara, H. (1983) *Nucleic Acids Res.* 11, 5037-5044.
6. Stam, J. C., Kwakman, J., Meijer, M. and Stuitje, A. R. (1986) *Nucleic Acids Res.* 14, 6871-6884.
7. Friefeld, B. R., Licky, J., Gronostajski, R. M., Guggenheimer, R. A., Krevolin, M. D., Nagata, K., Hurwitz, J. and Horwitz, M. S. (1986) *Curr. Top. Microbiol. Immunol.* 110, 221-255.
8. Stark, M. J. R., Mileham, A. J., Romanos, M. A. and Boyd, A. (1984) *Nucleic Acids Res.* 12, 6011-6030.
9. Sor, F. and Fukuhara, H. (1985) *Curr. Genet.* 9, 147-155.
10. Hishinuma, F., Nakamura, K., Hirai, K., Nishizawa, R., Gunge, N. and Maeda, T. (1984) *Nucleic Acid Res.* 12, 7581-7597.
11. Stark, M. J. R., and Boyd, A. (1986) *EMBO J.* 5, 1995-2002.
12. Tokunaga, M., Wada, N. and Hishinuma, F. (1987) *Nucleic Acid Res.* 15, 1031-1046.
13. Fukuhara, H. (1987) *Nucleic Acids Res.* 15, 10046.
14. Wesolowski, M., Dumazert, P. and Fukuhara, H. (1982) *Current Genet.* 5, 199-203.
15. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
16. Devereux, J., Haeberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* 12, 387-395.
17. Maundrell, K., Nurse, P., Schonholzer, F. and Schwein-gruber, E. M. (1985) *Gene* 39, 223-230.
18. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).
19. Feiberg, A.P. and Vogelstein, B. (1984) *Anal. Biochem.* 137, 266-267.
20. Murphy, S., Tripodi, M. and Melli, M. (1986) *Nucleic Acids Res.* 14, 9243-9260.
21. Clare, J. and Farabaugh, P. (1985) *Proc. Natl. Acad. Sci. USA* 82, 2829-2833.
22. Mellor, J. P., Fulton, S. M., Dobson, M. J., Wilson, W., Kingsman, S. M. and Kingsman, A. J. (1985) *Nature* 313, 243-246.
23. Argos, P., Tucker, A.D. and Philipson, L. (1986) *Virology* 149, 208-216.
24. Kuzmin, E. V. and Levchenko, I. V. (1987) *Nucleic Acids Res.* 15, 6758.
25. Salas, M. (1983) *Curr. Top. Microbiol. Immunol.* 109, 89-104.
26. Stillman, B. W., Tamanoi, F. and Mathews M. B. (1982) *Cell* 31, 613-623.

27. Lichy, J. H., Field, J., Horwitz, M. S. and Hurwitz, J. (1982) Proc. Natl. Acad. Sci. USA 79, 5225-5229.
28. Paillard, M., Sedeloff, R. R. and Levings, C. S. III (1986) EMBO J. 4, 1125-1128
29. Yoshikawa, H. and Ito, J. (1982) Gene 17, 323-335.
30. Quinn, J. P. and McGeoch, D. J. (1985) Nucleic Acid Res. 13, 8143-8163.
31. Earl, P.L., Jones, E.V. and Moss, B. (1986) Proc. Natl. Acad. Sci. USA 83, 3659-3663.