A transcriptional barrier to expression of cloned toxin genes of the linear plasmid kl of Kluyveromyces lactis: evidence that native kl has novel promoters

Michael A. Romanos and Alan Boyd

Leicester Biocentre, University of Leicester, Leicester, LE1 7RH, UK

Received May 20, 1988; Revised and Accepted July 15, 1988

ABSTRACT

The killer toxin of Kluyveromyces lactis consists of three polypeptides encoded by the linear plasmid kl. We re-introduced the entire kl sequence, cloned on a circular replicating plasmid, into K.lactis strains lacking kl, and found that the resulting transformants did not produce toxin. The barrier to expression was found to be transcriptional: the four transcripts of native kl were absent, and instead shorter, aberrant kl transcripts were made. We determined the precise initiation sites of the four transcripts of native kl: these had very short untranslated leaders and mapped about 14bp downstream of an "upstream conserved sequence" (UCS). It appears that kl has novel promoters which are inactive on circular plasmids which replicate in the nucleus. This is consistent with the suggestion that native kl resides in the cytoplasm.

INTRODUCTION

Certain "killer" strains of the lactose-utilising yeast Kluyveromyces lactis secrete a protein toxin which is able to inhibit the growth of sensitive strains of K.lactis or other yeasts (1). The toxin has been reported to cause arrest of sensitive cells at G1 and may act by inhibiting membrane-bound adenylate cyclase (2). The pure toxin is a glycoprotein complex consisting of three polypeptide subunits (3): α (99kD) β (30kD) and γ (27.5kD).

Production of the toxin and the corresponding immunity phenotype are related to the presence of two linear DNA plasmids kl (8.9 kb) and k2 (13.4 kb). These plasmids are present at a high copy number and are unusual in having a high (A+T)-content of 73%, inverted terminal repeat elements of about 200 bp and covalently attached terminal proteins (1,4,5). Several lines of evidence indicate that toxin production and immunity are conferred by kl. Strains cured by irradiation with ultra-violet light lack both plasmids or occasionally retain k2; both types of cured strains are non-killers and are sensitive to toxin (6). A spontaneous mutant strain, NK2/1, which contains k2 and a defective version of kl, kl-NK2 (5.9 kb), is
The map of kl shows the four open reading frames, left and right inverted repeats (IRL and IRR), and the deletion present in the variation kl-NK2, all drawn to scale.

a non-killer which retains immunity to toxin (7). Since strains containing kl alone have not been found, it is likely that kl is dependent on k.1 for its maintenance.

The entire nucleotide sequence of kl has been determined: the plasmid contains four large open reading frames (ORFs 1 to 4, Fig.1a), each of which is transcribed (8,9). Amino-terminal sequence analysis of the individual toxin polypeptides has demonstrated that all three are encoded by ORFs 2 and 4 (3). Specifically, ORF2 encodes a precursor of the α and β subunits, and ORF4 encodes the γ subunit. The deletion present in kl-NK2 has been mapped precisely and only ORF2 is disrupted (9).

We are interested in obtaining expression of cloned copies of individual kl ORFs. This would enable us to determine the function of each toxin subunit and of the proteins encoded by ORFs 1 and 3 and to study each of these further by in vitro mutagenesis. However, we have been unable to detect expression (i.e. toxin production) from K.lactis transformed with standard plasmids carrying cloned kl sequences. The data presented here...
Table 1 Yeast strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Plasmids</th>
</tr>
</thead>
<tbody>
<tr>
<td>K.lactis</td>
<td>a prototroph</td>
<td>[k2⁺ k1⁺]</td>
</tr>
<tr>
<td>SD11</td>
<td>a trpl lac4</td>
<td>[k2⁻ k1⁺]</td>
</tr>
<tr>
<td>SD801</td>
<td></td>
<td>[k2⁻ k1⁻]</td>
</tr>
<tr>
<td>SD802</td>
<td></td>
<td>[k2⁺ k1⁻]</td>
</tr>
<tr>
<td>NK2/1</td>
<td>a his-</td>
<td>[k2⁺ k1-NK2]</td>
</tr>
<tr>
<td>MRK1</td>
<td>a trpl lac4</td>
<td>[k2⁺ k1-NK2]</td>
</tr>
<tr>
<td>S.cerevisiae</td>
<td>a sst2 rme ade2</td>
<td></td>
</tr>
</tbody>
</table>

Demonstrate that the failure is at the level of transcription; indeed none of the kl ORFs is transcribed normally when present on a circular replicating vector in K.lactis. This may reflect the unusual structure of promoters in kl DNA and suggests the existence of a specialised transcription mechanism for native kl.

MATERIALS AND METHODS

Yeast strains, growth conditions and transformation

The K.lactis and S.cerevisiae strains used are listed in Table 1. The killer strain SD11 was kindly provided by C.Hollenberg, NK2/1 by H.Pukahara, and RC631 by R.K.Chan. SD801 and SD802 are plasmid-cured strains derived from SD11, obtained by screening survivors of UV irradiation for the presence of k1 or k2, using colony hybridisation.

Standard methods developed for S.cerevisiae were used for the maintenance, growth and transformation of K.lactis. Growth was at 30°C in YPD (1% yeast extract, 2% peptone, 2% glucose) or in minimal medium (0.67% yeast nitrogen base, 2% glucose, plus additional nutritional requirements) for selection of transformed cells. Where cells were grown for toxin production, the medium was buffered with 0.05M citrate-phosphate pH6.6. For selection of Trp⁺ transformants, 1% casamino acids was added to minimal medium, resulting in a relatively rich medium lacking tryptophan. K.lactis strains were transformed by the lithium acetate procedure (10).

Construction of strain MRK1 by spheroplast fusion

The strain MRK1 was derived by transfer of plasmids from NK2/1 to SD801 using polyethylene glycol-induced spheroplast fusion essentially as described in ref.11. After fusion, spheroplasts were regenerated in isotonic minimal medium containing tryptophan only, thus counterselecting the plasmid "donor" NK2/1. Following a round of purification on the same
medium the His\(^{+}\) colonies isolated were mainly haploid but included a small proportion of larger colonies which were either prototrophic (presumed SD801::NK2/1 a/a diploids) or Trp\(^{+}\) (presumed SD801 a/a diploids). Colonies containing the plasmids should be immune to killer toxin and were selected by growth on YPD plates containing toxin. The presence of linear plasmids in MPK1 was confirmed by agarose gel electrophoresis of spheroplasts lysed in situ (3).

For the production of crude killer toxin, IF01267 was grown to an \(A_{600}\) of 3.0 in buffered YPD. The supernatant was concentrated 40-fold by stirred cell ultrafiltration using an Amicon YM-100 membrane (100kD mol. wt. cut-off). The concentrate was cleared by centrifugation, sterilised by passage through a 0.2\(\mu\)m filter, diluted with 1 volume of sterile glycerol and stored at -20\(^{\circ}\)C. Toxin selection plates consisted of YPD agar with a 5\(\mu\)l overlay of YPD agar containing toxin (concentration 4X that in original supernatant, assuming 100% recovery).

**Killer toxin assay**

In initial exploratory experiments, plate assays were used in which killer yeast colonies grown on a background lawn of sensitive cells gave rise to a halo of growth inhibition (1,7). The activity of concentrated culture supernatants was assayed using antibiotic assay discs on similar plates.

For the detection of small amounts of toxin, a much more sensitive, quantitative assay was developed in which culture supernatants were tested for their inhibitory effect upon the growth of a sensitive strain in liquid medium. The strain to be tested for toxin production was grown to an \(A_{600}\) of 1-3 in buffered YPD or minimal medium plus casamino acids. The cells were removed by centrifugation, the supernatant was filter-sterilised, then diluted with 1 volume of sterile glycerol. To assay toxin activity, 20\(\mu\)l was added to 180\(\mu\)l of a toxin-sensitive strain (RC631) at a cell density of 5 \(\times\) 10\(^5\) cells/ml in buffered YPD in a microtitre well (NUNCION). Serial dilutions of the toxin, made in buffered YPD containing 50% glycerol, were tested to give a dose-response curve. The plate was incubated at 28\(^{\circ}\)C for 20 hours and then the growth of the tester strain was measured by determining the \(A_{600}\) after suitable dilution.

**Derivation of pUCW18kl**

*K. lactis* strain SD11 was transformed to tryptophan-independence using pUCW18 (a derivative of pUC18 containing the *S. cerevisiae* TRPl-ARS1 EcoRI fragment inserted at the NdeI site) linearised by digestion with BamHI.

7338
Colonies arose at a frequency of $10^3$ per µg of plasmid. Most of the transformants exhibited an unstable Trp$^+$ phenotype, consistent with the presence of replicating plasmids. Plasmids from ten of these were rescued in *E. coli*, using ampicillin selection, and analysed by restriction mapping. Six of the plasmids appeared to contain an insertion of the linear plasmid kl into the polylinker region of pUC18. Four of the plasmids were further characterised by determining the nucleotide sequences across both junctions between the input plasmid and the inserted DNA (using the method referred to in 12). This analysis revealed that the four plasmids were all slightly different from one another in the extent of the slight loss of pUCW18 DNA, but all four contained inserted DNA consistent with the presence of an entire intact copy of the kl plasmid. Thus for example the isolate which was subsequently propagated as pUCW18kl (Fig.1b) possessed the following junction sequences:

using universal primer  ...TAGAGGA[ACACATGAA...

using reverse primer  ...CGGCGAT[ACACATGAA...

The four plasmids analysed in detail were indistinguishable from one another in restriction digests, and yielded data consistent with the presence of an entire copy of kl: nevertheless the sequence analysis revealed that they had all arisen independently. This lends strong support to the notion that the plasmids contain a copy of the kl plasmid which is not rearranged in any way.

**DNA extraction and measurement of plasmid copy number**

DNA was extracted from *K. lactis* cells as described previously for *S. cerevisiae* (13) except that the lysate was treated with proteinase K to ensure efficient recovery of plasmids having covalently bound protein. Southern hybridisation analysis was used to determine the relative copy number of linear kl-NK2 versus circular pUCW18kl plasmids in transformed MRK1. Radiolabelled probes were used to detect diagnostic bands from each plasmid in restriction digests of total yeast DNA. Quantitation of the radioactive bands excised from the nitrocellulose filter was by scintillation spectrometry. Probes consisted of isolated DNA fragments labelled to a specific activity of $2 \times 10^9$ cpm/µg using random oligonucleotide primer extension from mixed hexanucleotides (Pharmacia) with α-[32P]-dCTP and Klenow fragment (14).
**RNA extraction and Northern blot analysis**

RNA was extracted from cells by glass bead breakage as described previously (15) except that the breaking buffer was 4M guanidinium thiocyanate, 0.1M sodium citrate (pH7.0), 0.1M 2-mercaptoethanol, 0.5% sodium lauryl sarcosinate. Total RNA was analysed by electrophoresis in 1.5% agarose gels containing MOPS buffer and 2.2M formaldehyde then transferred to GeneScreen membranes and hybridised as recommended by NEN. Probes were labelled as described above.

**Primer extension**

Oligonucleotides were synthesised according to ref.16. 15-mers were chosen that should be complementary to the transcripts of Kl ORFs 1, 2, 3 and 4: OL1 (TGAATCAAGTCGCC), OL2 (TGAAGAAGTTGGAG), OL3 (AGCTAGTGGCAGAG) and OL4 (TACGCCTGCAGG), respectively. These were 5'-32P-phosphorylated to a specific activity of approx 6 x 10^5 cpm/ng using γ-[32P]-ATP (Amersham) and T4 polynucleotide kinase. For the primer extension, annealing of 50μg total RNA to 10ng of labelled primer was carried out in 20μl of 50M Tris.HCl pH 8.3, 50mM NaCl, 8mM MgCl₂, 25mM DTT and 40μg RNasin (Promega Biotec) at 42°C for 1 hr. The mixture was made up to 50μl by addition of salts and buffer (to the same final concentration), 1mM GTP, dCTP, dGTP and dTTP, 25μg/ml actinomycin D and 20u AMV reverse transcriptase (Pharmacia), and incubated for a further 2 hr. The reaction was stopped by addition of 0.2% SDS and 50mM EDTA. RNA was digested by boiling for 5 min in the presence of 0.2M NaOH. Reactions were neutralised by addition of HCl, carrier tRNA was added, and the cDNA recovered by ethanol precipitation. The precipitate was redissolved in 3μl of buffer and denatured by boiling with dye-formamide for 3 min.

For direct comparison, the same oligonucleotide primers were used in sequencing reactions. Suitable DNA templates were made from recombinant M13 phage containing the relevant sections of Kl (8) and were sequenced by the "dideoxy" method (17). The primer extension products were analysed by electrophoresis alongside sequencing reaction products in 10% polyacrylamide gels containing 7.5M urea-TBE. The gels were fixed in 10% acetic acid, dried and placed against Kodak XAR-5 film at -70°C with intensifying screens for autoradiography.

**RESULTS**

**Expression in K.lactis of Kl sequences cloned in replicating vectors**

Our initial attempts to obtain expression of cloned Kl were made with a plasmid containing a K.lactis ARS sequence (18), the S.cerevisiae URA3
marker, and most of kl, including all of ORFs 1-4. This sequence was reconstructed from sub-clones previously used for sequence analysis (8). To see whether the cloned kl genes were expressed, the plasmid was introduced into a K. lactis uraA strain containing k2 alone or k2 and kl-NK2. Such transformants did not secrete amounts of toxin detectable using plate halo assays (data not shown); however we had found this type of assay to be very insensitive. In the experiments described here we used a highly sensitive microtitre well assay for toxin activity.

These initial experiments also suffered from the limitation that the native linear plasmids appear to be unstable in uraA strains during transformation and growth (unpublished observations). We therefore turned to the use of a K. lactis replicating plasmid carrying the S. cerevisiae TRPl marker which complements the K. lactis trpl mutation. The plasmid used, pUCW18kl, contains the S. cerevisiae TRPl-ARS1 EcoRI fragment (ARS1 is inactive in K. lactis, ref.18), and an entire unaltered copy of the k1 sequence (Fig.1b, see Materials and Methods for plasmid derivation). The phenomenon whereby such plasmids arise will be reported in greater detail elsewhere. The plasmid pUCW18kl can replicate in K. lactis due to the presence in k1 DNA of two regions (ARS elements) which confer this ability (19). The linear plasmids, k1 and k2, show structural similarities to adenoviral DNA and may replicate by a similar mechanism, involving priming of DNA synthesis by the terminal proteins (20). Internal ARS elements would therefore appear to be unnecessary for k1 replication and their occurrence is most likely to be fortuitous: such elements have been obtained from a number of diverse sources (see ref.21 for review). In plasmid k1, both ARS elements lie outside ORF2 and therefore should not interfere with its expression.

As the most critical test of ORF2 expression from pUCW18kl, we introduced the plasmid into a strain containing k2 and k1-NK2 (MRK1, a trpl lac4 [k2+ k1-NK2]). If the cloned ORF2 were expressed it should complement the ORF2 deletion in k1-NK2, leading to the production of toxin. The transformant MRK1/pUCW18kl was grown under selection in buffered minimal medium + casamino acids to A600 = 2 and the culture filtrate was assayed for killer toxin activity: no toxin activity was detected in microtitre well assays (Fig.2). As a positive control, we assayed toxin production in a culture of SD11 transformed with Kp2, a TRPl/K. lactis ARS plasmid based on YRpl7 (Fig.1c). This strain was grown under the same conditions to control for the effect upon toxin production of growth in minimal medium.
Figure 2 Microtitre well assay of toxin in culture supernatants from MRKI/pUCW18kl (●) and SD11/KRP2 (○). Serial three-fold dilutions of the supernatants were mixed with one volume of glycerol and then tested for their ability to inhibit the growth of S.cerevisiae strain NO631. Toxin activity is indicated by a lowering of the final OD<sub>600</sub> of the NO631 culture.

The results of the assay showed that SD11/KRP2 culture filtrates had toxin activity which was readily detectable at a 1 in 100 dilution (Fig.2). Since MRKI/pUCW18kl culture filtrates had no detectable activity, we concluded that there was no significant expression of ORF2 from pUCW18kl.

In order to eliminate the possibility that the lack of detectable expression was caused by pUCW18kl being unstable or present at a low number of copies per cell, we determined these parameters for the same cultures assayed for the production of toxin. Suitable dilutions of the cultures were plated on minimal agar + tryptophan. For the strain MRKI/pUCW18kl, 68% of the cells were Trp<sup>+</sup>, while for SD11/KRP2 the figure was 21%. The same cells were also used to prepare DNA for Southern hybridisation analysis to determine the average copy number of pUCW18kl relative to kl-NK2 in MRKI/pUCW18kl. Digestion of the DNA with PstI generates diagnostic bands for kl-NK2 (3.5 kb) and pUCW18kl (6.4 kb). Using a kl
probe having the same degree of overlap with both fragments (a 0.5 kb Clal-BglII fragment, coordinates 6944 to 7457, numbered as in ref.8), the radioactivity bound to these fragments was found to be in the ratio 3 to 1. This indicated that the average copy number ratio of kl-NK2 to pUCW18kl is 3:1. We might therefore have expected 30% of wild-type toxin levels in culture filtrates of MRK1/pUCW18kl if ORF2 on this plasmid were expressed at its normal levels. Since this level of toxin activity should be readily detectable, our conclusion that ORF2 fails to be expressed from cloned kl on a circular ARS vector appears to be valid.

Transcription of kl ORFs from pUCW18kl

In order to determine whether the failure of expression of ORF2 from pUCW18kl was at the level of transcription, RNA was prepared from MRK1/pUCW18kl cells and analysed by Northern blotting and hybridisation. As controls, we prepared RNA from SD11 [k2+ k1+], SD11/KR2p, and untransformed MRK1 [k2+ k1-NK2]. Using a probe specific for ORF2
Nucleic Acids Research

(coordinates 3210 to 6769), the normal ORF2 transcript of 3.5 kb was detected in RNA from both SD11 and SD11/KRp2 (Fig.3). As expected, untransformed MRK1 lacked this transcript but contained a 0.5 kb ORF2-specific transcript, presumably derived from the deleted version of ORF2 present in kl-NK2. No normal 3.5 kb ORF2 transcript was detected in MRK1/pUCW18kl either; instead, broad RNA bands migrating at about 0.9 kb and 0.5 kb were seen. The use of sub-fragments of ORF2 as probes indicated that the 0.5 kb band consisted of two transcripts: the kl-NK2 derived transcript present in MRK1 RNA, and a co-migrating major transcript originating from the region between EcoRI and BamHI sites of ORF2 (data not shown). In addition, the 0.9 kb transcript appeared to comprise several species of similar size which hybridised to different parts of ORF2. The results therefore indicate a complex pattern of aberrant transcripts derived from ORF2 in pUCW18kl. We have not attempted to identify the precise origins of these aberrant transcripts.

Since ORF4 encodes one of the subunits of the toxin, it was important to establish that this gene was transcribed normally from kl-NK2, otherwise complementation of kl-NK2 by pUCW18kl could not have restored toxin production. We therefore probed a Northern blot of MRK1 RNA with an ORF4-specific probe. An ORF4 transcript of the expected size (0.9 kb) and abundance was found to be present (Fig.3).

The abnormal transcription of ORF2 explains the failure of pUCW18kl to restore toxin production in MRK1 by complementing the ORF2 defect in kl-NK2. To examine the transcription of ORFs 1,3 and 4 from pUCW18kl we analysed RNA prepared from SD802 \([k_2^+ k_1^0]\) transformed with this plasmid and compared it to RNA from IFO1267 \([k_2^+ k_1^+], \) same killer phenotype and kl transcripts as SD11). In SD802/pUCW18kl the only possible source of kl-specific RNA is pUCW18kl. Probes specific for each ORF were hybridised to Northern blots (Fig.4). As in our previously reported results (8), we again failed to detect the kl ORF1 transcript; however, the transcripts of ORF2 (3.5 kb), ORF3 (1.4 kb) and ORF4 (0.9 kb) were detected in IFO1267 RNA with their respective probes. None of these was detected in SD802/pUCW18kl RNA; instead abundant smaller aberrant transcripts from ORFs 1,2, and 4 were seen. We therefore concluded that none of the ORFs of kl is transcribed normally when re-introduced on a circular replicating plasmid, but that they instead give rise to aberrant transcripts which are not produced by native kl.
Figure 4  Northern blot analysis of transcripts specific for each of the ORFs 1-4 in strains harbouring native or cloned kl. RNA from (a) wild-type killer strain IFO1267 or (b) SD802/pOCH18kl was hybridised to probes specific for each of the kl ORFs 1-4, corresponding to paired tracks 1-4 respectively. The positions of the normal transcripts of ORF2 (3.5kb), ORF3 (1.4kb), and ORF4 (0.9kb) are indicated by arrows. The ORF1 transcript is not detected in Northern blots. The coordinates of the probes used are: ORF1, 636-1552; ORF2, 3210-6769; ORF3, 6944-7457; and ORF4, 7991-8826.

Possible promoter elements in kl in relation to initiation of transcription

Our results indicate that transcription of native kl may be unusual in some way since kl promoters apparently do not function on circular replicating vectors. Previously, examination of the nucleotide sequence of kl led to the identification of an upstream conserved sequence (UCS), located approximately 30 bp upstream of ORFs 1, 2 and 3, which could be a promoter element (8). We proceeded to determine the initiation sites of the kl transcripts to see whether the information would shed any light on UCS function. In order to do this we used primer extension rather than SI mapping, since SI nuclease may generate artefactual heterogeneity (22). This problem is likely to be greatest in regions of high (A+T)-content, as was found for the LYS2 gene of Saccharomyces cerevisiae (23), and was therefore of particular relevance to our study since in kl DNA the regions immediately upstream of the four ORFs have local (A+T)-contents approaching 100%.

Synthetic oligonucleotides complementary to each ORF transcript were
prepared. These were used in primer extension reactions with total RNA extracted from IF01267 and also in sequencing reactions with the appropriate M13 templates. When separated by electrophoresis on a gel, the latter can provide size markers precise to within one nucleotide, since the extension products in the two types of reaction are the same.

Using primer extension, the transcripts from each ORF, including ORF1, were detected (Fig.5). The bands seen were kl-specific since they were not found when SD802 RNA [k2° kl0] was used (not shown). In each case, except for ORF4, multiple initiation points were found 2-13 bp upstream of the initiation codon, in regions of dyad symmetry. For the ORF4 transcript, the initiation points were about 20 bp upstream of the initiation codon, beyond the dyad symmetry in the region between ORFs 3 and 4. In all cases the initiating nucleotide was often a U residue. Our results are in broad agreement with those of Sor and Fukuhara (9) who used S1 mapping; this is despite the potential for spurious results arising from the use of S1, as mentioned above. However, our more precise data enabled us to establish a fixed distance of about 14bp from the 3' end of the UCSs of ORFs 1, 2 and 3 to the major initiation sites. We were then able to identify a sequence similar to a UCS about 14bp upstream of the major initiation sites of ORF4 transcripts.

DISCUSSION

We have found that when kl DNA is reintroduced into K.lactis in the form of DNA cloned into a circular replicating plasmid, the linear plasmid genes are not correctly transcribed. Thus the ORF2 gene carried upon the plasmid pUCW18kl fails to complement the ORF2 deletion present in a strain of

Figure 5 Primer extension analysis of transcripts from the native kl plasmid. (a) ORF-specific oligonucleotides were used for primer extension with RNA extracted from IF01267, or with the corresponding recombinant M13 DNA template in sequencing reactions. The products were separated by electrophoresis in 10% polyacrylamide gels containing urea, and autoradiographed. The primer extension products are shown on the left of the A,C,G, and T sequencing reactions for each ORF. For each transcript multiple initiation points were found and these could be mapped precisely using the sequence ladder. (b) The locations are shown of the initiation points in the region upstream of each ORF, drawn with the potential stem-loop structure. Major and minor start sites are represented by large and small asterisks respectively. The initiating ATG codons and the UCS elements are boxed. For ORF4 a potential partial UCS is boxed.
K. lactis harbouring the kl variant kl-NK2. Indeed none of the kl ORFs is transcribed normally; instead ORFs 1, 2 and 4 give rise to aberrant, shorter transcripts. In the case of ORF2 these aberrant transcripts derive from internal regions of the gene, suggesting that they are due to the existence of fortuitous promoters and terminators within ORF2, and that the normal promoter is inactive. In an analogous case, fortuitous yeast promoters are frequently found in bacterial DNA sequences, for example in the yeast vector YCpl9 (24). Our results indicate that the transcription of native kl is different from that of yeast chromosomal genes.

An examination of the regions upstream of the kl ORFs reveals significant differences from the equivalent portions of yeast chromosomal genes. Promoters for genes from S. cerevisiae are diverse in structure, but most extend over at least 100 bp and have a TATA box 50-100 bp upstream of the initiation site of transcription (25, 26). Genes from K. lactis appear to follow the same pattern (e.g., LAC4; ref.27). In kl the intergenic regions are much shorter, down to 11 bp, and a search for TATA boxes upstream of the ORFs is futile as the DNA is rich in A+T. There is however a sequence motif located upstream of ORFs 1, 2, and 3; we refer to this sequence, ACTA/TAATATATGA, as the UCS (8). In each case this UCS occurs at a similar position relative to the initiation codon of the ORF, -33/34 to -24/26 bp upstream (Fig.6). The conservation is particularly striking since in two cases the UCS lies partly (ORF2) or wholly (ORF3) within another ORF. We had previously suggested that the UCS may act as a promoter, although the lack of a similar sequence upstream of ORF4 appeared somewhat puzzling. For ORFs 1, 2 and 3 the region between the UCS and the initiating codon is almost entirely A+T, and contains a potential stem-loop structure with the initiation codon lying at its base (Fig.6). The region upstream of ORF4 has a similar potential stem-loop structure due to the symmetry of the 11 bp intergenic region of ORFs 3 and 4, and given this symmetry we had thought it conceivable that the ORF3 UCS might also function for ORF4 (8).

To examine the function of the promoter regions active in native kl, we determined the initiation sites of kl transcripts present in the RNA of a wild-type killer strain of K. lactis. The results show that each transcript has a very short 5' untranslated leader sequence (2-20 bases) with multiple initiation sites, and that the transcripts often start with a U residue. For chromosomal genes from S. cerevisiae, mRNAs most often initiate with an A residue and typically have untranslated leaders of 20-100 nucleotides.
Figure 6  Regions upstream of k1 ORFs 1 to 4 and k2 ORFs A and B. Dyad symmetry is indicated by arrows, and UCS elements or sequences bearing some similarity to these are boxed. ORF A of k2 would encode a basic polypeptide of 103 amino acids.

(e.g., CYC1, ref.28). These initiation sites are thought to be determined by the local sequence context rather than by their precise position relative to promoter elements (26). The initiation sites for the ORF1, 2 and 3 RNAs are centred about 14 bp downstream of the UCS and are distributed across potential stem-loop structures in a manner which suggests that they are determined by position rather than by the nature of the initiating nucleotide. We had not previously identified a UCS for ORF4, and its transcript is different from the others in that it does not initiate in a potential stem-loop structure. If the UCS of ORF3 could function for ORF4 also, it might be expected that both transcripts would initiate divergently in the same stem-loop. Another possibility is that ORF4 has its own upstream element with only partial identity to the consensus UCS, and therefore hitherto undetected. We therefore searched the appropriate region, about 14 bp upstream of the major initiation sites of ORF4 RNA, for sequences similar to a UCS. The sequence TAAAATATCTGA with partial identity to a UCS (matches underlined) was found in this precise location (Fig.6), providing support for this view.

Due to our present inability to modify native k1 we cannot test the importance of the UCS elements and the stem-loops in k1 promoters. However, the natural mutant plasmid k1-NK2 may shed light upon the function of the latter. In k1-NK2 a deletion starting in the loop of the ORF2 stem-loop structure removes most of ORF2. The resulting truncated ORF2 is transcribed efficiently to give a 0.5kb RNA (see Fig.3) though the original potential stem-loop structure is eliminated. We are currently mapping the 5'-end of this transcript in order to see whether another stem-loop has
been recruited, or if the transcript starts at the usual distance of 14 bp downstream of the UCS.

The observations discussed above strongly suggest that kl has novel promoters, which are recognised by a specific transcription factor. Given the indication from plasmid curing experiments that kl is dependent upon k2 for replication and/or maintenance, this factor could be k2-coded, in which case we would expect k2 genes to have similar promoter structures. An examination of the published terminal sequence data for k2 (29) reveals the presence of ORFs at either end. Strikingly, in each case the initiating ATG codon is preceded in the anticipated position by a sequence resembling the UCS identified in kl (Fig. 6). These sequences also have some similarity to the partial UCS upstream of ORF4 in kl. It must be emphasised, however, that nothing is known about the transcription of k2.

There may be some support for the notion that kl is dependent on k2 for transcription. Firstly, homology searches using the polypeptide encoded by kl ORF1 have shown significant similarities to family B DNA polymerases (30, 31, 32). This, taken with evidence suggesting that ORF3 is involved in the immunity phenotype (33), would leave no space on kl to encode a transcription factor. Secondly, recent nucleotide sequence data for k2 has led to the identification of an ORF encoding a potential polypeptide of 982 residues which exhibits significant similarities to certain DNA-dependent RNA polymerases of bacterial origin (Duncan Wilson, pers. comm.). Therefore k2 may encode an RNA polymerase specific for kl and k2.

We have thought of two explanations, which are not mutually exclusive, for the existence of a transcriptional barrier to the expression of kl ORFs from circular recombinant plasmids. Firstly, the kl promoters may only be in an active conformation in linear DNA. Secondly, and we believe more likely, native kl and k2 may reside in the cytoplasm and be transcribed by an enzyme present only in the cytoplasm, whereas circular plasmids are most likely sequestered in the nucleus. There is indeed some circumstantial evidence to support the view that the linear plasmids are cytoplasmic. Thus when they are introduced in S. cerevisiae the plasmids are only stable in \( p^0 \) strains (which lack all mitochondrial (mt) DNA; ref.34); this incompatibility may reflect a competition for a specific extranuclear DNA replication factor. Furthermore, when the linear plasmids are transferred to a \( p^0 \) strain, fluorescent particles appear in the cytoplasm after staining with DAPI (35). Another line of evidence relates to the base composition and nucleotide sequence of kl. In known cytoplasmic DNAs such
as the mtDNA of *S. cerevisiae* or *K. lactis*, drift towards a high A+T content is thought to result from an inaccessibility of the DNA to uracil excision-repair mechanisms present in the nucleus (36). The sequences of mtDNA genes are also found to have a relatively low information content (37). Plasmid Kl resembles mtDNA by both of these criteria (8). Finally, the ease of release of linear plasmid DNA from gently lysed cells has been taken as evidence for a cytoplasmic location (38). The linear plasmids are unlikely to be located within the mitochondria since they confer the killer phenotype upon $\rho^0$ strains of *S. cerevisiae*, where there is no mitochondrial protein synthesis (36).

All of the available evidence points then towards the linear plasmids possessing a pronounced degree of autonomy in replication and gene expression. This autonomy is an intriguing feature of the biology of these plasmids, but also creates significant problems for rapid analysis of plasmid function. In the case of ORF3 it appears that the expression barrier is not insurmountable since this gene has been found to be weakly expressed from circular plasmids in *K. lactis*, at levels sufficient to confer some degree of immunity to toxin (33). It is striking that ORF3 is the only one of the four genes of Kl which we find does not generate aberrant transcripts. Our own preliminary attempts to direct expression of ORF2 using the yeast ADH1 promoter in *K. lactis* and *S. cerevisiae*, have failed so far (unpublished data); this could be due to the presence of 'nuclear' terminators within ORF2. It may be that the further analysis of gene function for the linear plasmids will depend upon developing the ability to manipulate and reintroduce linear Kl into *K. lactis* cells.

**ACKNOWLEDGEMENT**

We thank John Keyte for synthesising oligonucleotides and Jackie Swallow for excellent technical assistance. The plasmid KRp2 was kindly provided by Duncan Wilson.

Present addresses: $^+$Wellcome Biotech, Langley Court, Beckenham, Kent BR3 3BS and $^*$Department of Biochemistry, University of Edinburgh Medical School, Edinburgh, EH8 9XD, UK

**REFERENCES**