Volume 12 Number 19 1984 Nucleic Acids Research

Cloning and nucleotide sequences of the linear DNA killer plasmids from yeast

Fumio Hishinuma*, Kenzo Nakamura+, Keiko Hirai, Rei Nishizawa, Norio Gunge and Tadakazu Maeda

Mitsubishi-Kasei Institute of Life Sciences, 11 Minamiooya, Machida-shi, Tokyo 194, Japan

Received 17 April 1984; Revised and Accepted 10 September 1984

ABSTRACT

The linear DNA killer plasmids (pGKL1 and pGKL2) isolated from a Kluyveromyces lactis killer strain are also maintained and expressed its killer character in Saccharomyces cerevisiae. After these killer plasmid DNAs isolated from S. cerevisiae were treated with alkali, four terminal fragments from each plasmid DNAs were cloned separately. Using these and other cloned DNA fragments, the terminal nucleotide sequences of pGKL2 and the complete nucleotide sequence of pGKL1 were determined. The inverted terminal repetitions of 202 bp and 182 bp were found in pGKL1 and pGKL2, respectively. The pGKL1 sequence showed an extremely high A + T content of 73.2% and it contained five large open reading frames. The largest of these open reading frame was suggested to code for a membrane-bound precursor of glycoprotein subunit of the killer toxin.

INTRODUCTION

The yeast Kluyveromyces lactis IFO1267 contains two linear DNA plasmids called pGKL1 (8.9 kb) and pGKL2 (13.4 kb) (1). K. lactis harboring both plasmids produces a killer toxin that kills a certain group of yeast including Saccharomyces cerevisiae, Saccharomyces italicus, Saccharomyces rouxii, Kluyveromyces lactis, Kluyveromyces thermotolerans, Kluyveromyces vanudenii, Torulopsis gabrate, Candida utilis and Candida intermedia (1). It has been shown, by curing and deletion mapping of pGKL1, that the killer toxin gene and the immunity-determining gene reside on pGKL1, and pGKL2 is essential for the maintenance of pGKL1 since pGKL1 has never been found when pGKL2 was lost (2). These pGKL plasmids were transfered from a K. lactis killer strain into a nonkiller strain of S. cerevisiae by protoplast fusion and transformation. S. cerevisiae harboring both plasmids showed the same killer phenotype as the donor K. lactis (3, 4).

The killer toxin coded by pGKL1 is quite different from the well characterized K1 killer toxin which is coded by S. cerevisiae double-

stranded linear RNA plasmid, called M1-dsRNA (5). While the K1 killer toxin act as a protonophore causing the leakage of ATP in sensitive cells (5, 6), the pGKL killer toxin specifically inhibits adenylate cyclase activity and causes G1 arrest in sensitive cells (7). In contrast to the K1 killer toxin which is a small protein of 9.5 and 9.0 kilodaltons (kd) (5, 8), the pGKL killer toxin is a large glycoprotein with a molecular weight larger than 100 kd which can be dissociated into subunit polypeptides (9, Fujimura et al. manuscript in preparation). The k1 killer toxin has been reported to be secreted from the cell by proteolytic cleavage of a larger membrane-bound glycoprotein precursor (5). The mechanism of secretion of the pGKL killer toxin is not known.

In order to determine the primary structure and the gene organization in pGKL1 and pGKL2, we have cloned the entire nucleotide sequences of both plasmids into E. coli. Sequencing analysis of the cloned pGKL plasmids revealed that pGKL1 and pGKL2 have terminal inverted repeat sequences of 202 and 182 bp, respectively, and that pGKL1 sequence has five open reading frames larger than 100 amino acids.

MATERIALS AND METHODS

Strains, media, enzymes, and chemicals

S. cerevisiae F102-2 (a, leu2-3, 2-112, his4-519, can1, pGKL1, pGKL2) that was a fusant between K. lactis 2105-1D (α, ade1, ade2, leu, pGKL1, pGKL2) and S. cerevisiae AH22 (a, leu2-3, 2-112, his4-519, can1), a sensitive strain against the pGKL killer toxin, was grown in YEPD medium and used to isolate pGKL plasmids as previously described (1). E. coli K-12 strains C600 (F^-, thi-1, thr-1, leuB6, lacY1, tonA21, supE44) and JA221 (F^-, $hsdM^+$, $hsdR^-$, leuB6, lacY, ΔtrpE5, recA1 were used as host for the cloning of DNA fragments of pGKL plasmids.

Bacterial alkaline phosphatase, T4 DNA ligase, and polynucleotide kinase were purchased from Bethesda Research Laboratories (BRL). Terminal deoxynucleotidyl transferase was from P. L. Biochemicals (PL), Klenow fragment of E. coli DNA polymerase I was from Boehringer-Mannheim (BM), and Zymolyase 60000 was from Seikagaku Kogyo (SK). The restriction endonucleases were purchased from BRL, PL, BM, SK, Takara Shuzo, and Nippon Gene. The enzymes were used as recommended by the

manufacturer. α-^{32}P-dNTPs, γ-^{32}P-ATP, and α-^{32}P-3'dATP were from Amersham.

Cloning of DNA fragments of pGKL plasmids in E. coli

The cleared lysate from S. cerevisiae F102-2 was prepared by using Zymolyase as described previously (1), treated with 0.1N NaOH for 30 min at 37°C, and neutralized with 0.1N HCl. The alkali treated lysate was extracted three times with 0.1M Tris-HCl (pH 8.0) saturated phenol, and twice with ether. DNA in the aqueous phase of the final extraction was precipitated with ethanol, redissolved in TE (10mM Tris-HCl, pH7.4, 1mM EDTA), and subjected to Sephacryl S-1000 column chromatography in order to remove small RNAs. The pGKL1 and pGKL2 eluted with chromosomal DNA were digested with BamHI which cuts each plasmid at a single site, and ligated with BamHI and SmaI digested pLS354 which is a promoter-probing shuttle vector in E. coli and B. subtilis consisting pBR322, pE194 and pUB110 (10). Ligated DNA was introduced to E. coli C600 and recombinant plasmids containing pGKL1 and pGKL2 were identified from ampicillin resistant transformants by colony hybridization. Probes were made from the purified pGKL1 and pGKL2. The plasmid DNAs were isolated from the hybridization-positive colonies and analyzed for their physical maps. Four different recombinant DNAs were identified and designated, respectively, as pGKF106 (containing the right half of pGKL1), pGKF107 (the left half of pGKL1), pGKF201 (the right half of pGKL2) and pGKF202 (the left half of pGKL2).

The plasmid pGKF219 was constructed as follows; the alkali treated pGKL2 as described above was digested with XbaI and the termini of the DNA fragments were filled in to make blunt ends with Klenow fragment of DNA polymerase I and 4 deoxynucleoside triphosphates. DNA fragments with blunt ends were inserted in a SmaI site of pUC8 (11). The pGKF101 is a recombinant plasmid containing the 6.4 kb PstI fragment of pGKL1 inserted into PstI site of pBR322. The restriction enzyme maps of the plasmid DNAs used here are shown in Fig. 1.

DNA sequencing

DNA fragments were labeled at the 5' ends with γ-^{32}P-ATP and polynucleotide kinase, and at the 3' end with α-^{32}P-dNTP and Klenow fragment of DNA polymerase I or α-^{32}P-3'dATP and terminal deoxynucleotidyl transferase. Nucleotide sequences were determined by the chemical method of Maxam and Gilbert (12). The entire regions reported

here were sequenced repeatedly (more than twice) for each strands.

Preparation of RNA fractions and Northern blotting

RNA was prepared from log-phase yeast cells essentially by the method of Struhl and Davis (13), except for 22 μg of Zymolyase 60000/ original ml of culture was used for spheroplasting cells. The total RNA samples were twice precipitated by 2M LiCl at 4°C for 10-16 hr (14), and loaded on an oligo-dT cellulose (Collaborative Research) column equilibrated with 0.5M LiCl, 10mM EDTA, 0.2% SDS, 10mM Tris-HCl (pH 7.5) (15). PolyA$^+$ RNA was eluted from the column by 1mM EDTA, 10mM Tris-HCl (pH 7.5), and precipitated with 2.5 volumes of 5M LiCl-:ethanol (1:24, by vol.). Oligo-dT cellulose column chromatography was repeated once more.

RNA samples were denatured in 50% dimethylsulfoxide, 1M glyoxal, 10mM sodium phosphate (pH 7.0) at 50°C for 1 hr (16), and separated by gel electrophoresis in a horizontal 1% agarose. Transfer of RNA from the gel to diazotized paper, pretreatment, hybridization and washing of diazotized paper were carried out as described by Alwine et al. (17). DNA fragments used as ^{32}P-probe were prepared by nick translation method of Rigby et al. (18).

RESULTS

Cloning of pGKL plasmid DNAs

Two major types of terminal structures in linear DNA species are known; one has a hairpin structure and the other has a terminal protein at 5' end of DNA. If pGKL plasmids have terminal protein(s) at the 5' ends, they should be cloned in E. coli only after removal of the terminal protein(s). It has been known that the proteins covalently attached to linear DNAs are successfully removed by the treatment with 0.1M NaOH for 30 min at 37°C (19, 20). Since only one BamHI cleavage site exists in either pGKL1 or pGKL2 (Fig. 1), pGKL plasmid DNAs were digested with BamHl and ligated to BamHI and SmaI cleaved pLS354. E. coli C600 was transformed with ligated DNA and the transformants carrying pGKL fragments were screened from ampicillin resistant transformants by colony hybridization using ^{32}P-labelled pGKL1 or pGKL2 as probes. Seventy-five hybridization positive clones were isolated from about 1200 ampicillin resistant colonies tested. The plasmid DNAs were isolated from 18 positive clones, and finally four different recombinant DNAs that contain each of the left and

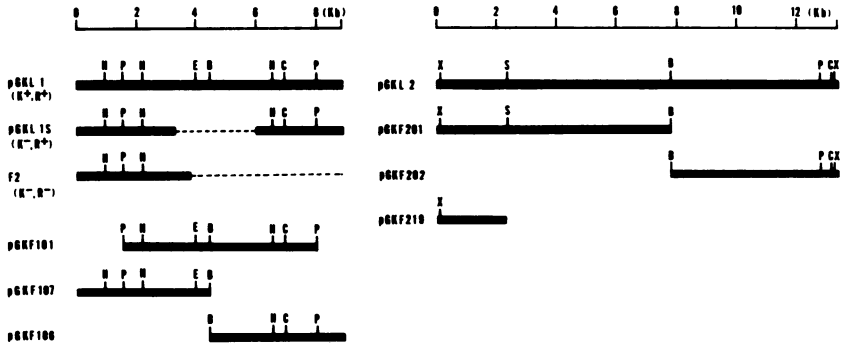


Fig. 1 Structures of the linear killer DNA plasmids and the recombinant plasmids containing the fragments of the linear plasmid DNAs. The linear killer DNA plasmids, pGKL1 abd pGKL2, were originally isolated from K. lactis IFO1267. The deletion plasmids of pGKL1, pGKL1S and F2, were isolatd from K. lactis NK$^-$-1 and S. cerevisiae Pd11-8, respectively. The other plasmids are the recombinant DNAs containing the DNA fragments of pGKL1 and pGKL2 as indicated in the figure. The recombinant plasmids (pGKF106, pGKF107, pGKF201, pGKF202, and pGKF219) were constructed by the use of pUC8 as a vector, and pGKF101 was cloned a HindIII fragment in a HindIII site of pBR322. The vectors are not shown in this figure. The following abbreviations are used for the restriction endonucleases. H: HindIII, P: PstI, N: NdeI, E: EcoRI, B: BamHI, C: ClaI, X: XhoI, S: SacI

right halves of pGKL1 or pGKL2 were obtained. Plasmid pGKF106 and pGKF107 contain the right and the left half of pGKL1, respectively, and pGKF201 and pGKF202 contain the right and the left half of pGKL2, respectively. The physical maps of these plasmid DNAs are shown in Fig. 1.

Although we do not know exactly whether the termini of the pGKL plasmids are blunt ends or 5' (or 3') protruding ends, we conclude from the success to clone the termini in SmaI site as mentioned above that they are blunt ends.

Nucleotide Sequences of the terminal regions of pGKL plasmids

In order to determine nucleotide sequences of pGKL1 and pGKL2, the fine restriction maps were made (Fig 2a). These maps suggested that pGKL plasmids might have inverted terminal repetitions (ITR) since AluI-XhoI-TaqI-HinfI sites in pGKL2 or two RsaI sites in pGKL1 exist at same distance in the left and the right termini of each plasmid. The nucleotide sequences determined by Maxam and Gilbert method (12) confirmed this, that is, terminal 202 bp of the left and right termini of pGKL1 were entirely identical (Fig. 3), and 182 bp

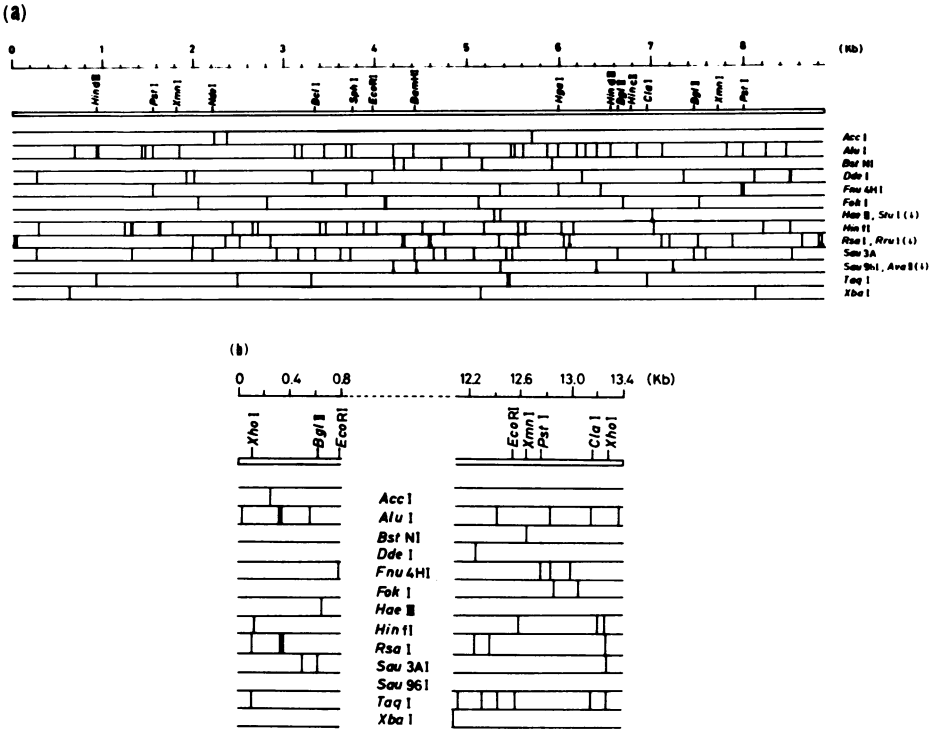
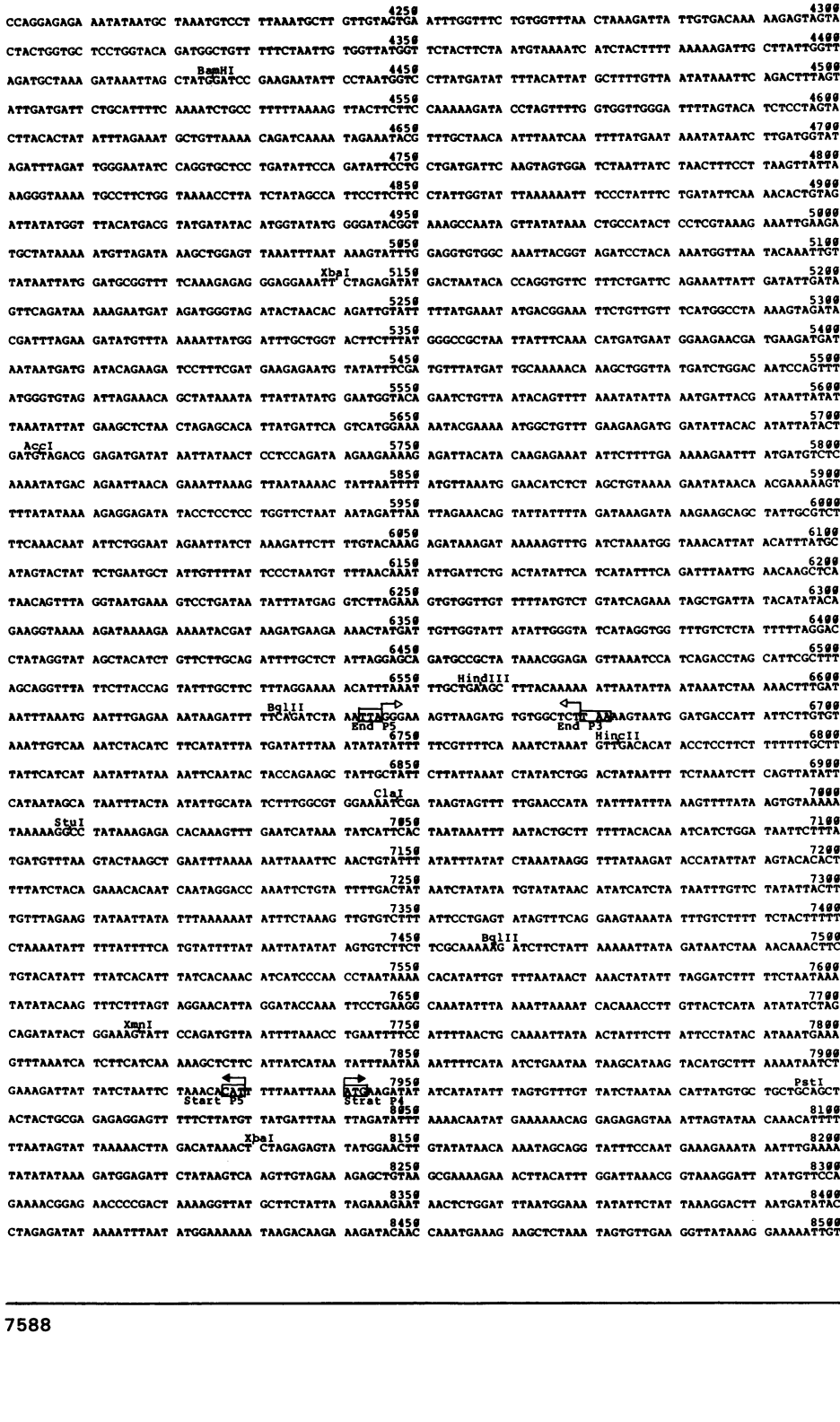


Fig. 2 The restriction enzyme maps of the entire sequence of pGKL1 (a) and the terminal regions of pGKL2 (b). The maps were obtained by analysis of cloned pGKL1 and pGKL2 fragments.

of both terminal sequences of pGKL2 showed also inverted repetition (Fig. 4) except that the sequence of the right side of pGKL2 was 2 bp ($\frac{AA}{TT}$) longer than that of the left side. To examine whether this 2 bp difference comes from an artifact during cloning or whether the left side is truly 2 bp shorter than the right side in pGKL2, we isolated independently a recombinant plasmid (pGKF219) containing the left terminal region of pGKL2 using pUC8 (11) as a vector. Nucleotide sequence of pGKF219 showed that the terminal sequence begins with $\frac{AAAGG}{TTTCC}$. This sequence is still one bp shorter than that of the right (pGKF201). Thus, the number of A-T base pairs of the left terminus in pGKL2 could not be settled in this report. If it is 4 A-T base pairs as the right terminus, the ITR sequence of pGKL2 shall be 184 bp. The conclusion must wait until the nucleotide sequence of the native pGKL2 is determined.


Start P1
XbaI
HindIII
PstI
XmnI
End P1
Start P2
NdeI
AccI
AccI
End P2
Start P3
EcoRI


4250
4300
4350
4400
BamHI
4450
4500
4550
4600
4650
4700
4750
4800
4850
4900
4950
5000
5050
5100
XbaI
5150
5200
5250
5300
5350
5400
5450
5500
5550
5600
5650
5700
AccI
5750
5800
5850
5900
5950
6000
6050
6100
6150
6200
6250
6300
6350
6400
6450
6500
HindIII
6550
6600
BglII
6650
6700
End P5
End P3
HincII
6750
6800
6850
6900
ClaI
6950
7000
StuI
7050
7100
7150
7200
7250
7300
7350
7400
7450
BglII
7500
7550
7600
7650
7700
XmnI
7750
7800
7850
7900
PstI
7950
Start P5
Strat P4
8000
8050
8100
XbaI
8150
8200
8250
8300
8350
8400
8450
8500

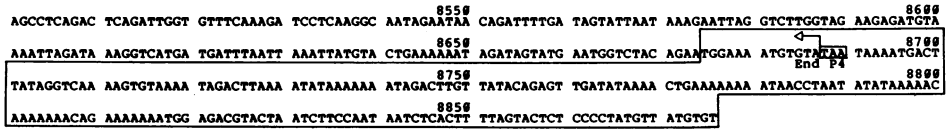


Fig. 3 Nucleotide sequence of the entire pGKL1. The cleavage sites for restriction endonucleases, 2Ø2 bp of the ITR (large boxes), and the location of start and end of five open reading frames (small boxes and arrows) are indicated.

The ITR of both pGKL1 and pGKL2 are AT rich : 76.2% for pGKL1 and 69.8% for pGKL2, and contain many A or T clusters. The clusters more than 4 consecutive A or T stretches are 12 in ITR of pGKL1 and 9 in ITR of pGKL2. No apparent homology was found between the ITR sequences of pGKL1 and pGKL2 by a computer analysis (21).

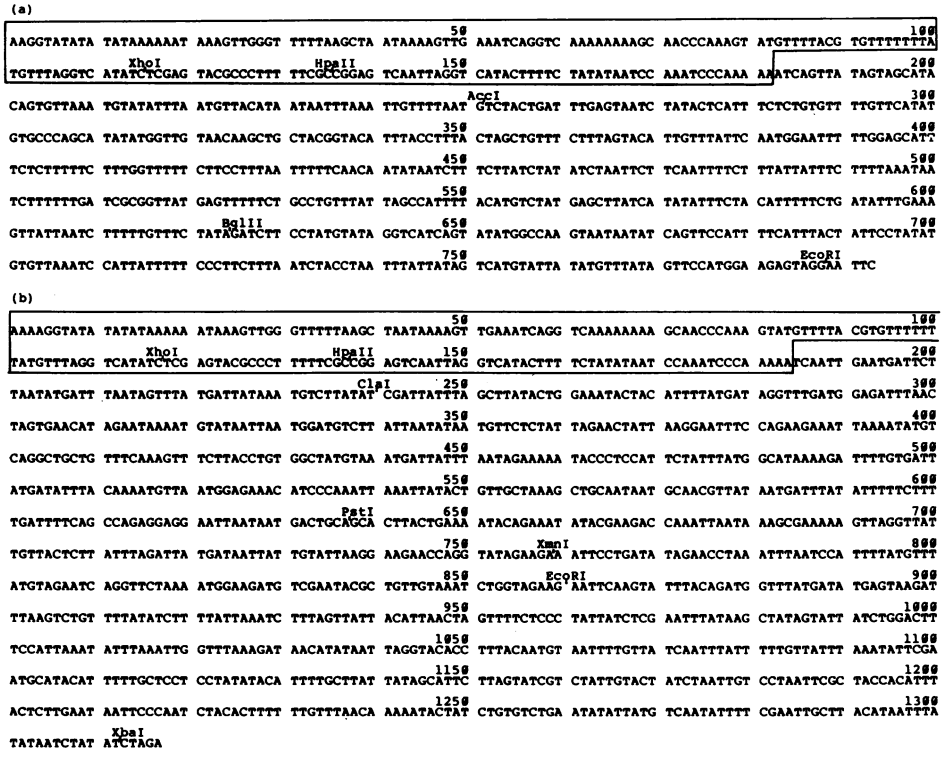


Fig. 4 Nucleotide sequence of the terminal regions of pGKL2. The restriction enzyme cleavage sites and the ITR sequence regions are indicated. (a) The left terminal of pGKL2. (b) The right terminal of pGKL2.

Fig. 5 Translation stop codons in pGKL1. The positions of trinucleotide sequences (TAA, TAG, and TGA) are marked in 6 different frames. Arrows show the direction of translation.

Sequence organization in pGKL1

The killer toxin and the resistance or immunity to the killer is coded by pGKL1, while pGKL2 is required for the maintenance of pGKL1 in a cell (2, 3, 4). Two deletion mutants of pGKL1 which affect the killer phenotype have been obtained in our laboratory, and their restriction enzyme maps are shown in Fig. 1. Deletion mutant pGKL1S is defective in killer secretion, but retains resistance to the killer (2), while F2 has lost not only the ability to secrete killer but also the resistance to the killer (4). Details of deletion mapping of pGKL1S, F2 and its tail-to-tail dimer, F1 (4), will be described elsewhere. These deletion mapping data indicate that a structural gene for the killer and the resistance gene locates at its middle and the right portion of pGKL1, respectively, shown in Fig. 1.

In order to analyze the gene organization in pGKL1, the entire nucleotide sequence of pGKL1 was determined by the use of the cloned DNA fragments of pGKL1 (Fig. 3). The number of total nucleotides of pGKL1 was 8876 bp and its A + T content was 73.2% which is consistent with the low buoyant density of pGKL1 DNA (1).

When we analyzed the termination codons (TAA, TAG TGA) of translation in 6 different frames, five open reading frames (P1 to P5) being able to code for proteins with more than 100 amino acids were found in pGKL1 (Fig. 5, Table 1). All of these open reading frames were terminated with TAA codon. Since we could not find in pGKL1 the consensus sequence for splicing in yeast (TACTAAC) (22), these regions may be translated independently if they are transcribed.

Among the five open reading frames, only P3 overlapped with both of the sequences deleted in pGKL1S and F2 defective in killer secretion. Especially, a 2.7 kb deletion of pGKL1S located entirely inside

TABLE 1 : The open reding frames and the characteristics of the predicted proteins from them in pGKL1

	Nucleotide No.	A + T (%)	No.of Amino Acids	Molecular Weight
P1	213-1973	75.0	586	68579
P2	2048-3202	72.4	384	45042
P3	3232-6672	70.3	1146	128720
P4	7941-8690	72.7	249	28709
P5	7929-6643	77.5	428	50873

The open reading frames in pGKL1 were analyzed in six different frames. The results revealed that 5 protein coding regions (P1-P5) were expected in pGKL1. The number of amino acids and the molecular weight of each possible protein were calculated by computer analysis using MRC program (42). It was assumed that the protein is translated from the first ATG codon of each reading frame.

of P3. The location of ATG initiation codons within P3 indicate that P3 can code for a polypeptide with the maximum number of amino acid residues of 1146 (Table 1). These results suggest that P3 codes for killer toxin polypeptide(s).

In order to examine whether the P3 region of pGKL1 is actually expressed in yeast cells, we analyzed RNA transcripts by Northern blotting. A 2.6 kb EcoRI-HindIII fragment of pGKL1 (from 3973 to 6561) covering most of the P3 (Fig. 2a) hybridized to a poly A^+ mRNA from S. cerevisiae F102-2 carring pGKL plasmids (Fig. 6, lanes 7-9), which migrated to the same position as 25S rRNA (3360 $\pm$ 80 bases; 23). We could not detect any hybridization of the same probe neither to total and poly A^- RNA from the same strain (Fig. 6, lanes 1-6), nor to poly A^+ mRNA from S. cerevisiae AH22 lacking pGKL plasmids (data not shown). A 660 bp DdeI fragment located adjacent to the left side of the EcoRI site (from 3307 to 3971) also hybridized to a S. cerevisiae F102-2 poly A^+ mRNA with the same electrophoretic mobility as 25S rRNA (data not shown). These results indicate that almost the entire region of P3 is actually transcribed into a mRNA of about 3.4 kb in S. cerevisiae F102-2.

DISCUSSION

After treatment of the pGKL plasmids with alkali, the four fragments of the plasmid DNAs containing the each terminal were cloned.

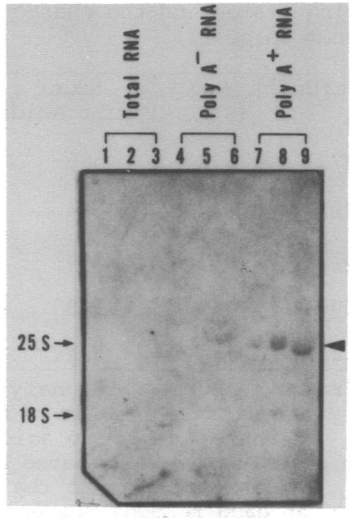


Fig. 6 Northern blott analysis of mRNA transcribed from the P3 portion of pGKL1. Total RNA, poly A^- RNA and poly A^+ RNA fractions from S. cerevisiae F102-2 carring pGKL plasmids were analyzed by Northern blotting with 2.6 kb EcoRI-HindIII fragment of pGKL1 as a ^{32}P-probe. The position of 25S and 18S rRNAs are indicated by thin arrows, and the position of RNA hybridized with the probe is indicated by thick arrow. The amount of RNA applied to each lane are, 1: 5 µg, 2: 10 µg, 3: 20 µg, 4: 5 µg, 5: 10 µg, 6: 20 µg, 7: 2 µg, 8: 4 µg, 9: 8 µg.

Since the treatment of the plasmid DNAs with 0.1M NaOH at 37°C for 30 min does not probably affect polynucleotide linkages, one can assume that the termini of the pGKL plasmid DNAs are blunt ends, and is associated with alkali labile material(s) at their 5' ends.

Some of the linear DNA species other than chromosomes have been found and characterized for their terminal structures in order to understand the mechanisms of DNA replication. There are two major types of linear DNA species. The first type posseses the terminal protein at 5' ends of DNA like adenovirus DNA (24), phages ø29, ø15, M2Y, and N_f from Bacillus subtilis (19, 20, 25) and the linear plasmids from Streptomyces rochei (pSLA1 and pSLA2) (10). All these DNAs have been characterized to carry ITR sequences. The second type has a hairpin structure, like rDNA of Tetrahymena (26, 27) that contains 20-70 repeats of the hexanucleotide CCCCAA at or near the DNA terminus. Similar structures have been found in rDNA molecule of Physarum in which the repeating unit is CCCTA (28). The telomeres of S. cerevi-

siae chromosomes seem to be similar in structure to rDNA of Tetrahymena (20, 30).

The nucleotide sequences of pGKL1 and pGKL2 clearly revealed that these plasmid DNAs have the ITR sequences. These structures are very similar to the termini of adenovirus DNAs where 102 to 162 bp of the ITR sequences have been found and the terminal protein is associated at their 5' ends (31-34). The yield of pGKL plasmid DNA was drastically reduced when the cleared lysate from S. cerevisiae F102-2 was extracted with phenol. These observations suggest that some protein is associated to the pGKL plasmid DNAs. Recently, we have detected peptide(s) associated at or near their 5' ends (35). This indicates that the DNA replication in pGKL plasmids might be initiated by a protein priming mechanism as reported for adenovirus DNA. It is, however, noteworthy to point out that the consensus sequences for the initiation origin of DNA replication (ars) in S. cerevisiae ($^{A}_{T}TTTAT^{A}_{G}TTT^{A}_{T}$; 36) are found once (139-149) in the ITR sequences of both pGKL plasmids. Then, it is interesting to test whether the ITR sequences have an activity of the initiation of DNA replication in S. cerevisiae and K. lactis.

In the course of preparation of this paper, Sor et al. (37) reported the sequence of the terminal 227 bp of pGKL1 and pGKL2 determined with the native plasmid DNAs isolated from K. lactis. Since 5' ends of the plasmid DNAs are not labeled by kination, they deduced their sequence of 5' terminus from that of the opposite strands. Their sequences, however, were identical to our results, but they suggested that 5' ends of the plasmid DNAs might be protruding. We do not know the reason why their prediction is different from ours, and it may be explained by the difference of origins of plasmid DNAs analyzed. In our case, the pGKL plasmids were prepared from S. cerevisiae F102-2 which is a fusant between K. lactis 2105-1D and S. cerevisiae AH22.

The right end of pGKL2 begins with AAAAGG, whereas the left terminus begins with AAGG in pGKF202 and AAAGG in pGKF219. The difference in the number of A-T base pairs might reflect the state of this plasmid DNA in S. cerevisiae, or come from an artifact during cloning in E. coli. Since Sor et al. (37) reported that the both termini of pGKL2 begin with AAAAGG, it might be possible that the terminal structure of the pGKL plasmids are not exactly identical in

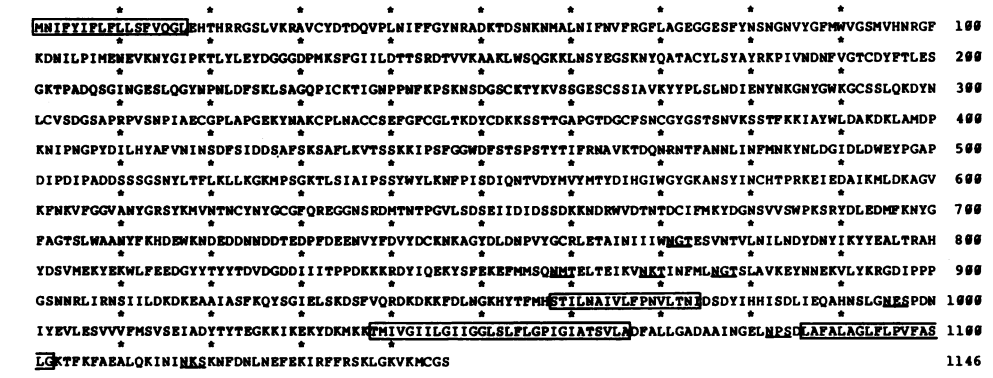


Fig. 7. Amino acid sequence of a polypeptide which can be coded by P3. The amino acid sequence starting from the first ATG codon is deduced from the DNA sequence. The location of hydrophobic segments which is longer than 17 amino acid residues, with more than 65% hydrophobic amino acids and without charged amino acids are indicated by boxes. Underlined sequences represent the potential sites for Asn-linked glycosylation.

these two host cells.

At least the genes for killer toxin and immunity-determinant reside on pGKL1 (2, 3, 4). The pGKL killer toxin is quite large, and it can be resolved into three protein subunits by SDS-polyacryamide gel electrophoresis (9, Fujimura et al., manuscript in preparation). The largest subunit is a glycoprotein, and its apparent molecular weight varies between 97 and 123 kd depending on conditions of gel electrophoresis. The molecular weight of this subunit decreses about 14 kd upon deglycosylation with endoglycosidase H. The other two subunits are simple proteins with apparent molecular weight of 25 and 24 kd (Fujimura et al., manuscript in preparation). The molecular nature of the immunity-determinant is not known.

When the protein coding capacity of pGKL1 was analysed by searching the termination codons (TAA, TAG, TGA) in 6 different frames, 5 proteins of larger than 100 amino acids were predicted to be coded by pGKL1 (Fig. 6 and Table 1). Several lines of evidence suggest that P3 codes for a precursor protein of at least the glycoprotein subunit of the killer toxin. First, only P3 overlapps with both of the sequences deleted in pGKL1S and F2 defective in killer secretion. Especially, a 2.7 kb deletion of pGKL1S lies inside of P3. Second, almost the entire region of P3 is actually transcribed into a poly A^+ mRNA (Fig. 6). Third, among the five open reading frames, only P3 can code for a

protein large enough to account for the polypeptide chain of glycoprotein subunit (Table 1). Fourth, the amino acid sequence of a protein which can be coded by P3 (Fig. 7) shows characteristic features of precursors for exocellular glycoproteins. If we assume that the translation starts from the first ATG codon, the amino-terminal segment of 17 amino acids is highly hydrophobic (Fig. 7) and resembles to signal peptides found in various secretory protein precursors, which directs polypeptides for processing and secretory pathways. It lacks, however, the amino-terminal basic amino acid(s) which is found in many, but not all, signal peptides (38). It is noticed that the pentapeptide sequence ^{7}Phe-Leu-Phe-Leu-Leu^{11} found in this segment is also present in the signal peptide of pre-invertase of S. cerevisiae (39, 40), and pre-invertase signal peptide also lacks amino-terminal basic amino acid (40). The amino acid sequence of a protein coded by P3 also contains seven potential asparagine-linked major glycosylation site sequences (Asn-X-Thr/Ser; Fig. 7).

The size of a putative precursor coded by P3 is considerably larger than the polypeptide chain of glycoprotein subunit, even if the removal of signal peptide is taken into account. In addition to the amino-terminal signal peptide-like sequence, the carboxy-terminal portion of putative precursor contains three hydrophobic segments without charged amino acids (Fig. 7). Two of these segments located closer to carboxy-terminus are especially hydrophobic in terms of their length or hydrophobic amino acid content. We have recently detected a membrane-bound protein larger than the glycoprotein subunit which was immunologically cross-reactive with anti-killer toxin antiserum in S. cerevisiae F102-2. This membrane protein was absent in killer-minus cells. Furthermore, the size of this membrane protein decreased about 12 kd when cells were labeled in the presence of tunicamycin which inhibited the secretion of killer toxin (Fujimura, et al., manuscript in preparation). It seems likely that the secretion of glycoprotein subunit of pGKL killer toxin into the culture medium involves the proteolytic cleavage of membrane-bound larger precursor. Such a proteolytic cleavage of membrane-bound precursor has been shown to occur in the secretion of 9.5 and 9.0 kd killer toxin coded by S. cerevisiae M1-dsRNA plasmid (5, 8) and secretory component (SC) of human epithelial cells (41). Substantiation of this model, however, requires the determination of amino-terminal sequence

of the glycoprotein subunit and the comparison of peptide maps of the glycoprotein subunit, the membrane-bound precursor and the *in vitro* translation product of mRNA transcribed from pGKL1.

The open reading frames P4 or P5 could be the immunity-determinant since the deletion mutant F2 lacking about 5 kb of the right side of pGKL1 (Fig. 1) does not afford resistance to the killer.

ACKNOWLEDGEMENT

We are grateful to Drs. M. Kageyama and Y. Kikuchi for discussion and critical reading of the manuscript, to S. Ohtsuka for the computer analysis of the nucleotide sequence, and to T. Mukohara for help in the preparation of the manuscript.

*To whom correspondence should be addressed
+Present address: Laboratory of Biochemistry, Faculty of Agriculture, Nagoya University, Chikusa-ku, Nagoya 464, Japan

REFERENCES

1. Gunge, N., Tamaru, A., Ozawa, F. and Sakaguchi, K. (1981) J. Bacteriol. 145, 382-390.
2. Niwa, O., Sakaguchi, K. and Gunge, N. (1981) J. Bacteriol. 148, 988-990.
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. Tipper, D.J. and Bostian, K.A. (1984) Microbiol. Rev. 48, 125-156.
6. de la Peña, P., Barros, F., Gascon, S., Lazo, P.S. and Ramos, S. (1981) J. Biol. Chem. 256, 10420-10425.
7. Sugisaki, Y., Gunge, N., Sakaguchi, K., Yamasaki, M. and Tamura, G. (1983) Nature 304, 464-466.
8. Bostian, K., Bussey, H., Elliot, Q., Burn, V., Smith, A. and Tipper, D.J. (1984) Cell 36, 741-751.
9. Sugisaki, Y., Gunge, N., Sakaguchi, K., Yamasaki, M. and Tamura, G. (1984) Eur. J. Biochem. 141, 241-245.
10. Hirochika, H., Nakamura, K. and Sakaguchi, K. (1984) EMBO J. 3, 761-766
11. Vieira, J. and Messing, J. (1982) Gene 19, 259-268.
12. Maxam, A. and Gilbert, W. (1980) In "Methods in Enzymology, Vol. 65", Grossman, L. and Moldave, K. Eds., pp. 499-560, Academic Press, New York.
13. Struhl, K and Davis, R.W. (1981) J. Mol. Biol. 152, 535-552.
14. McAlister, L. and Finkelstein, D.B. (1980) J. Bacteriol. 143, 603-612.
15. Finkelstein, D.B., Strausberg, S. and McAlister, L. (1982) J. Biol. Chem. 257, 8405-8411.
16. Carmichael, G.G. and McMaster, G.K. (1980) In "Methods in Enzymology, Vol. 65", Grossman, L. and Moldave, K. Eds., pp. 380-391, Academic Press, New York.
17. Alwine, J.C., Kemp, D.J., Parker, B.A., Reiser, J., Renart, J.,

Stark, G.R. and Wahl, G.M. (1979) In Methods in Enzymology, Vol. 68", Wu, R., Ed., pp. 220-242, Academic Press, New York.
18. Rigby, P.W.J., Dickmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
19. Escarmis, C. and Salas, M. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 1446-1450.
20. Salas, M., Mellado, R. P., Vinuela, E., and Sogo, J. M. (1978) J. Mol. Biol. 119, 269-291.
21. Korn, L. J., Queen, C. L. and Wegman, M. N. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 4401-4405.
22. Langford, C. and Gallwitz, D. (1983) Cell 33, 519-527.
23. Philippsen, P., Thomas, M., Kramer, R.A. and Davis, R.W. (1978) J. Mol. Biol. 123, 387-404.
24. Robinson, I. and Padmanabhan, R. (1980) Biochem. Biophys. Res. Commun. 94, 398-405.
25. Yoshikawa, H. and Ito, J. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 2596-2600.
26. Blackburn, E. H. and Gall, J. G. (1978) J. Mol. Biol. 120, 33-53.
27. Yao, M. -C. and Yao, C. -H. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 7436-7439.
28. Bergold, P. J., Campbell, G. R., Littau, V. C. and Johnson, E. M. (1983) Cell 32, 1287-1299.
29. Murray, A. W. and Szostak, J. W. (1983) Nature 305 189-193.
30. Szostak, J. W. and Blackburn, E. H. (1982) Cell 29, 245-255.
31. Alestrom, P., Stenlund, A., Li, P. and Pettersson, U. (1982) Gene 18, 193-197.
32. Steenbergh, P. H., Maat, J., Van Ormondt, H. and Sussenbach, J. S. (1977) Nucleic Acids Res. 4, 4371-4389.
33. Stillman, B. W., Topp, W. C. and Engler, J. A. (1982) J. Virol. 44, 530-537.
34. Tolun, A., Alestrom, P. and Pettersson, U. (1979) Cell 17, 705-713.
35. Kikuchi, Y., Hirai, K. and Hishinuma, F. (1984) Nucleic Acids Res. 12, 5685-5692.
36. Broach, J. R. Li, Y. Y., Feldman, J., Jayaram, M., Abraham, J., Nasmyth, K. A. and Hicks, J. B. (1983) Cold Spring Harb. Symp. Quant. Biol. 48, 1165-1173.
37. Sor. F., Wesolowski, M. and Fukuhara, H. (1983) Nucleic Acids Res. 11, 5037-5044.
38. Inouye, M. and Halegoua, H.O. and Cannon, L.E. (1980) CRC Crit. Rev. Bioche. 7. 339-371.
39. Perlman, D., Halvorson, H.O. and Cannon, L.E. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 781-785.
40. Carlson, M., Taussing, R., Kustu, S. and Botstein, D. (1983) Mol. Cell. Biol. 3, 439-447.
41. Mostov, K.E. and Blobel, G. (1982) J. Biol. Chem. 257, 11816-11821.
42. Staden, R. (1977) Nucleic Acids Res. 4, 4037-4051.

Note Added in Proof

During the preparation of this revised manuscript, we noticed that M. J. R. Stark et al. reported the nucleotide sequence of pGKL1 (K1) appeared in Nucleic Acids Res., 12, 6011 (1984). The total number of K1 was determined by them to be 8874 bp which is 2 bp shorter than ours. From the consequence in this 2 bp difference, they deduced that the open reading frames of P1 and P2 were contiguous in one large open reading frame.