Saccharomyces cerevisiae contains two discrete genes coding for the α-factor pheromone

Arjun Singh, Ellson Y.Chen, June M.Lugovoy, Chung N.Chang, Ronald A.Hitzeman and Peter H.Seeburg

Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Blvd., South San Francisco, CA 94080, USA

Received 21 March 1983; Revised and Accepted 13 May 1983

ABSTRACT

Two genes, MFα1 and MFα2, coding for the α-factor in yeast Saccharomyces cerevisiae were identified by in situ colony hybridization of synthetic probes to a yeast genomic library. The probes were designed on the basis of the known amino acid sequence of the tridecapeptide α-pheromone. The nucleotide sequence revealed that the two genes, though similar in their overall structure, differ from each other in several striking ways. MFα1 gene contains 4 copies of the coding sequence for the α-factor, which are separated by 24 nucleotides encoding the octapeptide Lys-Arg-Glu-Ala-Glu(or Asp)-Ala-Glu-Ala. The first α-factor coding block is preceded by a sequence for the hexapeptide Lys-Arg-Glu-Ala-Glu-Ala and 83 additional amino acids. MFα2 gene contains coding sequences for two copies of the α-factor that differ from each other and from α-factor encoded by MFα1 gene by a Gln → Asn and a Lys → Arg substitution. The first copy of the α-factor is preceded by a sequence coding for 87 amino acids which ends with Lys-Arg-Glu-Ala-Val-Ala-Asp-Ala. The coding blocks of the two copies of the pheromone are separated by the sequence for Lys-Arg-Glu-Ala-Asn-Ala-Asp-Ala. Thus, the α-factor can be derived from 2 different precursor proteins of 165 and 120 amino acids containing, respectively, 4 and 2 copies of the pheromone.

INTRODUCTION

Sexual conjugation between cells of opposite mating types of the yeast Saccharomyces cerevisiae involves diffusible factors produced by the **a** and α cells (1,2,3). The factor produced by the α haploid cells, variously referred to as α-factor, α-pheromone, α mating pheromone, or α-factor pheromone, has been thoroughly characterized both physiologically and chemically. The α-factor is produced constitutively by α cells and is secreted in the culture medium (1,2). The α-pheromone induces a morphological change in **a** cells (2) characteristic of the mating process, and reversibly arrests the growth of these cells in the early G1 period of the cell cycle (2,4). It has no such effects on α or **a**/α cells. It has been suggested that arrest of **a** cells by α-factor at a specific stage is a method of synchronization of the two haploid cell cycles prior to mating, thus ensuring that only unbudded and mononucleate haploids fuse to form diploid zygotes

(4,5). The finding that many mating-defective mutants do not produce the pheromone (6,7) further supports the important role of the α-factor in the mating process.

Duntze and coworkers (8,9) first determined that the α-factor is a family of four oligopeptides of 12-13 amino acid residues having the basic sequence H_2N-(Trp)-His-Trp-Leu-Gln-Leu-Lys-Pro-Gly-Gln-Pro-Met (or MetSO)-Tyr-COOH. This was soon confirmed by a number of different laboratories (10,11,12). Experiments with temperature-sensitive mutants and with cycloheximide indicate that the pheromone is synthesized by translation of specific mRNA (13). Studies by Scherer et al. (13) and recently published experiments by Julius et al. (14) indicate that the α-pheromone is produced by processing of a larger precursor. In this report we describe the molecular cloning and the nucleotide sequence of two α-factor genes, MFα1 and MFα2, and their flanking regions. The genes were isolated using synthetic oligonucleotide probes which were based on the known amino acid sequence of the pheromone. While these studies were in progress, Kurjan and Herskowitz (15 and personal communication) have isolated and characterized a yeast pheromone gene that is nearly identical to one of the pheromone genes, MFα1, described in this report. The cloning strategy used by these authors was based on the phenotype of a yeast mutant defective in α-factor secretion (15).

MATERIALS AND METHODS

Bacterial and Yeast Strains. E. coli K-12 strain 294 (endA thi$^-$ hsr$^-$ hsm$^+$) (16) was used for bacterial transformations. Yeast strain A147 (a leu2) was constructed using standard yeast genetics procedures (17). Yeast strain 483 (a leu2 his4 can1) is from our culture collection. The yeast strain 20B-12 (α trp1 pep4) was obtained from the Yeast Genetics Stock Center, University of California, Berkeley.

Growth Media. The routine yeast growth medium contained 1 percent Bacto-yeast extract, 2 percent Bacto-peptone and 2 percent dextrose. Yeast minimal medium contained 0.67 percent Bacto-yeast nitrogen base without amino acids, 2 percent dextrose and 3 percent agar. The minimal medium supplemented with 1M sorbitol was used for yeast transformations. Bacterial growth medium was LB (18) which was supplemented with 20 μg/ml ampicillin when used for transformation. S-agar plates used for colony screening contained per liter: 32g tryptone, 5g NaCl, 15g Difco agar and 0.2g NaOH to which ampicillin or chloramphenicol was added as indicated.

Transformations. E. coli 294 was transformed using a published procedure

(19). Yeast were transformed essentially as described (20,21).

Enzymes and DNA Preparations. Restriction enzymes were purchased from New England Biolabs and Bethesda Research Laboratories and were used according to manufacturer's recommendations. T4 DNA ligase was from New England Biolabs and was used in 20mM Tris-HCl (pH 7.5), 10mM $MgCl_2$, 10mM dithiothreitol, 1mM ATP at 14°. Calf alkaline phosphatase was purchased from Boehringer Mannheim and was used in 100mM NaCl, 50mM Tris-HCl (pH 7.4), 10mM $MgSO_4$, 1mM 2-mercaptoethanol at 37°.

Plasmid DNAs were prepared by the cleared lysate method (22) and were purified by Bio-Rad Agarose A-50 column chromatography. Small amounts of plasmid DNAs from individual E. coli transformants were prepared by a quick-screening procedure (23). DNA restriction fragments were isolated by electroelution from a 1 percent agarose gel followed by phenol/chloroform extraction and ethanol precipitation. Oligo-deoxynucleotide probes prepared by the phosphotriester method (24) were kindly provided by Dr. R. Crea.

Design of the Probes. The 15-mer oligonucleotide probes for the α-factor gene were designed on the basis of the amino acid sequence of the pheromone (9) and yeast codon usage frequencies. The rationale is outlined in Fig. 1 where the last 5 amino acids of the α-factor and all the possible codons and their usage frequencies are given. The codon usage is derived from the total of 2 different glyceraldehyde-3-phosphate dehydrogenase clones (25,26) and of alcohol dehydrogenase I (B. Hall, personal communication). The codon usage for these and other genes has recently been summarized (27). As can be seen from Fig. 1, virtually all of the most likely sequences coding for the 5 amino acids are included in the nucleotide sequence 5'-GG$\frac{T}{C}$CAACC$\frac{A}{T}$ATGTAC. Accordingly, two pools consisting of two oligonucleotides each, and complementary to above sequence, were synthesized (see Fig. 1). No other contiguous 5 amino acids in the pheromone could be covered with such a limited set of oligonucleotides.

Screening of Recombinant Plasmids. A genomic library made by insertion of partially Sau3A-digested yeast DNA into the BamHI site of YRp7 (28) was screened for presence of α-factor gene clones. E. coli transformants were grown on nitrocellulose filter paper (Schleicher and Schuell, BA85) placed on S-agar plates containing 20 μg/ml ampicillin. After 6 hours at 37°, filters were transferred to S-agar plates containing 150 μg/ml chloramphenicol. After 15 hours of amplification colonies were tested for hybridization using a modified in situ colony screening procedure (29). ^{32}P-labelled (30) synthetic oligonucleotides were used as hybridization

Carboxy terminus of α-factor:	Gly	--	Gln	--	Pro	--	Met	--	Tyr COOH
Possible codons and their usage	GGU (90) GGC (3) GGA (0) GGG (0)		CAA (20) CAG (0)		CCA (32) CCU (3) CCC (1) CCG (0)		AUG (20)		UAC (33) UAU (0)
Consensus oligonucleotides:	5'-GG$^{T}_{C}$CAACC$^{A}_{T}$ATGTAC								
Synthesized oligonucleotide pools complementary to above:	I. 5'-GTACATTGGTTG$^{A}_{G}$CC II. 5'-GTACATAGGTTG$^{A}_{G}$CC								

Figure 1. Design of the α-factor probes. The codon usage frequencies indicated in parentheses are for three yeast genes described in the text.

probes. Filters were hybridized overnight at 42° in 10mM Tris (pH 7.5), 6mM EDTA, 0.1mM ATP, 1mM sodium pyrophosphate, 0.8M NaCl, 1X Denhardt's solution, 0.5 percent NP-40, and 0.1 mg/ml E. coli tRNA. Filters were washed 3 times for 20 min in 6XSSC at 30°. Dried filters were exposed to Kodak XR-2 X-ray film with DuPont Lightning-Plus intensifying screen at -80°.

DNA Sequence Determination. DNA sequence determination was as previously described (31). Briefly, DNA sequences were obtained by the chain termination method (32) using recombinant phages M13 mp8 and mp9 (33) as the source for single-stranded "template" DNA and a synthetic oligonucleotide for priming E. coli DNA polymerase I (large fragment, Boehringer Mannheim) in the presence of dideoxynucleoside triphosphates using α-^{32}P dCTP (400 Ci/mmole, Amersham) for labeling synthesized DNA chains. Reactions were electrophoresed on 5 percent polyacrylamide/8M urea "thin" gels (34). Gels were dried onto 3MM paper (Whatman) and exposed to X-ray film for 2 to 12 hr.

RESULTS

Identification of Recombinant Plasmids Containing the α-factor Gene. Approximately 4500 bacterial colonies containing recombinant plasmids were tested for in situ hybridization (29) with ^{32}P-end-labelled oligonucleotide pool I (Fig. 1). Twenty-four colonies hybridized to varying degrees. Small amounts of plasmid DNAs were prepared from these 24 colonies and tested for hybridization with the same probes after spotting the DNA samples on a nitrocellulose filter. Two of the 24 plasmids, designated as p51 and p52 respectively, hybridized strongly and were chosen for further study. The p51 and p52 plasmids also hybridized with the oligonucleotide pool II.

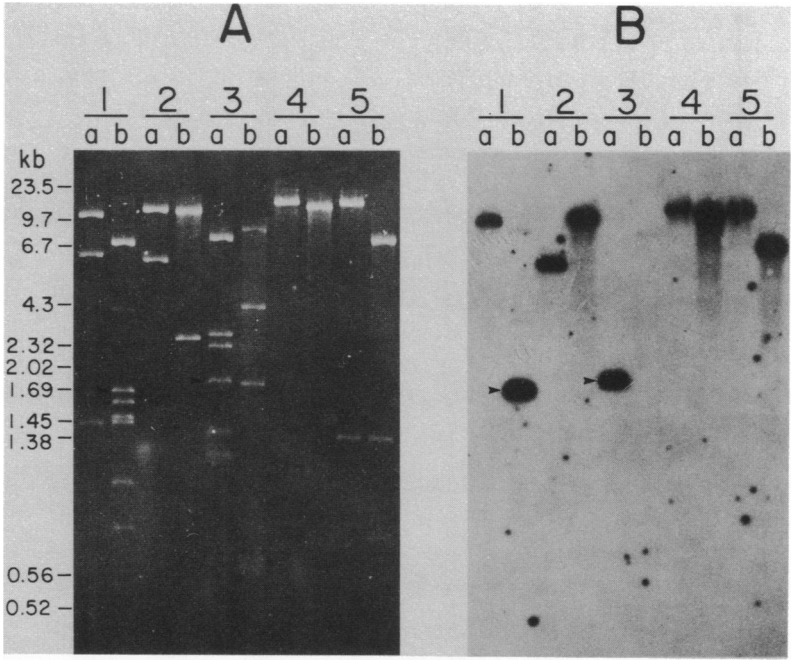

Figure 2. Localization of homology between the α-factor probes and the DNA fragments from p51 (a) and p52 (b) recombinant plasmids. The two plasmids were digested with different restriction endonucleases and then electrophoresed on a one percent agarose gel. The DNA fragments were transferred to nitrocellulose paper and hybridized to ^{32}P-labelled probes. Panel A: Ethidium bromide stained gel. Panel B: Southern blot. Lanes: 1, EcoRI; 2, SalI; 3, HindIII; 4, BamHI; 5 PstI. The arrows indicate the two DNA fragments that were subcloned. The size standards were derived from lambda, YRp7 or pBR322 DNA.

Subcloning of the Hybridizing Sequences. To characterize the inserts that hybridized with the synthetic probes, plasmid DNA prepared from the p51 and p52 clones was subjected to restriction enzyme analysis with EcoRI, SalI, HindIII, BamHI, and PstI. As seen in Fig. 2A, the 2 recombinant plasmids are quite dissimilar. Only EcoRI and PstI digestions of the two plasmids yielded one common fragment each. In both cases the common fragments arise from the YRp7 vector (28). The common 1.45 kbp EcoRI fragment is the TRP1 insert and the 1.38 kbp PstI piece is the DNA between PstI sites in the TRP1 and the amp^R genes.

We identified the fragments that contained sequences complementary to

the probe by the method of Southern (35). Fig. 2B shows that, except in one case, digestion with all 5 restriction enzymes yielded a fragment that specifically hybridized with the probe. No hybridization was seen with any of the fragments produced by restriction of p52 DNA with HindIII.

The smallest restriction fragments that contained sequences complementary to the synthetic probes were the 1.7 kbp EcoRI fragment from p52 and the 1.8 kbp HindIII fragment from p51. These two DNA fragments were isolated from a preparative agarose gel by electroelution and separately ligated to appropriately cleaved plasmid pBR322 (36) DNA. The ligation mixture was used to transform E. coli 294 and the plasmid DNA from transformants was analyzed by a quick-screen procedure (23). Two transformants, designated p53 and p56, containing the 1.7 kbp EcoRI and 1.8 kbp HindIII fragment inserts, respectively, were analyzed as follows: Plasmid DNA was prepared from p53 and p56 and digested separately with BamHI, ClaI, PvuI, PstI, and SalI. The resulting DNA fragments were separated on a 1 percent agarose gel, transferred to nitrocellulose filter paper (35) and tested for hybridization with ^{32}P-labelled probes. The analysis of the restriction digests and corresponding hybridization patterns of the p53 DNA, the recombinant plasmid containing the 1.7 kbp yeast DNA as an EcoRI fragment, showed that the yeast DNA in this clone contained one SalI site and two PstI sites and that the sequence complementary to the probes was included within a 0.5 kbp PstI-SalI fragment. The HindIII fragment of yeast DNA in clone p56 lacked recognition sites for these enzymes, and the linearized plasmid, resulting from cleavage at single recognition sites for these enzymes in the pBR322 vector, hybridized with the probes. This plasmid was then digested with a number of additional restriction endonucleases and the digests were analyzed by the method of Southern as described above. We found that the hybridizing sequences in this plasmid were contained on a 1.3 kbp HindIII-SacI fragment.

Structure of the α-Factor Genes. The nucleotide sequences of large parts of the 1.7 kbp EcoRI fragment and the 1.3 kbp HindIII-SacI fragment are shown in Fig. 3 and Fig. 4, respectively. The p53 sequence contains an open reading frame coding for a protein of 165 amino acid residues which carries 4 internal repeat units within its C-terminal half. Each unit begins with Lys-Arg and ends with the α-factor sequence. Within each unit the pair of basic residues is separated from the α-factor by several Glu(or Asp)-Ala dipeptide repeats. The N-terminal half of the protein starts with a highly hydrophobic sequence of 22 amino acids which probably represents a signal sequence for secretion. As indicated in Fig. 3, the 61 amino acid

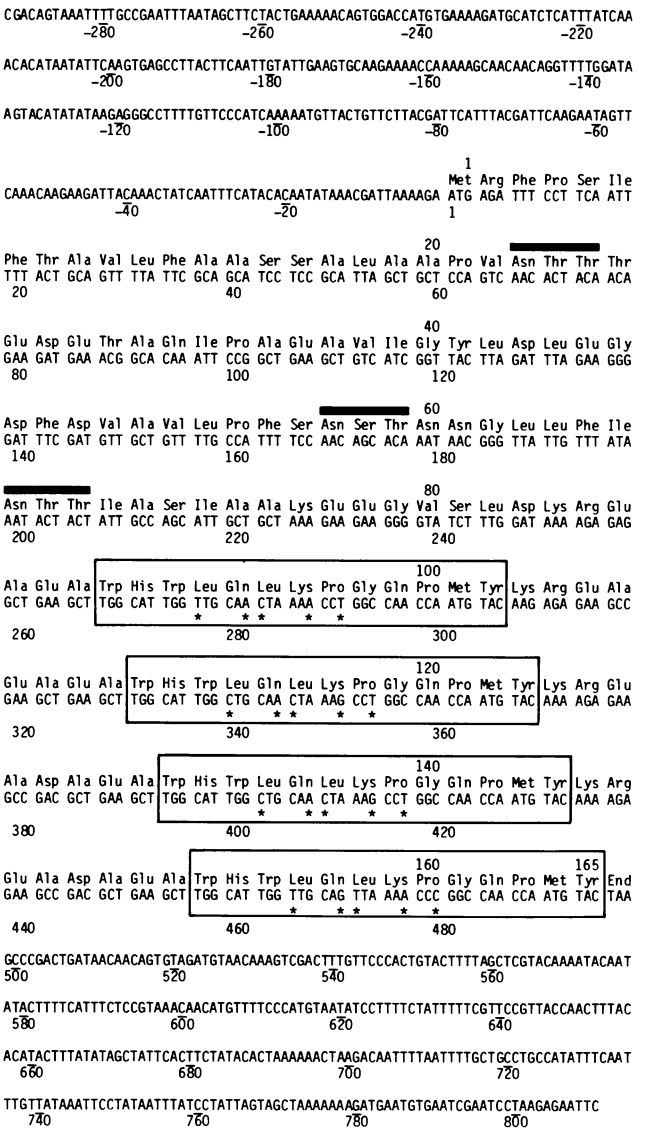

Figure 3. Nucleotide sequence of MFα1 gene and its nontranslated 5' and 3' flanking regions. The predicted amino acid sequence of the pheromone precursor is also shown. The numbers above and below the sequence denote the positions of amino acids and nucleotides, respectively. The four copies of the α-factor sequences are included in the boxed areas. The asterisks indicate differences in the nucleotide sequence in one or more copies of the α-factor coding regions. Three potential N-glycosylation recognition sites are indicated by bars.

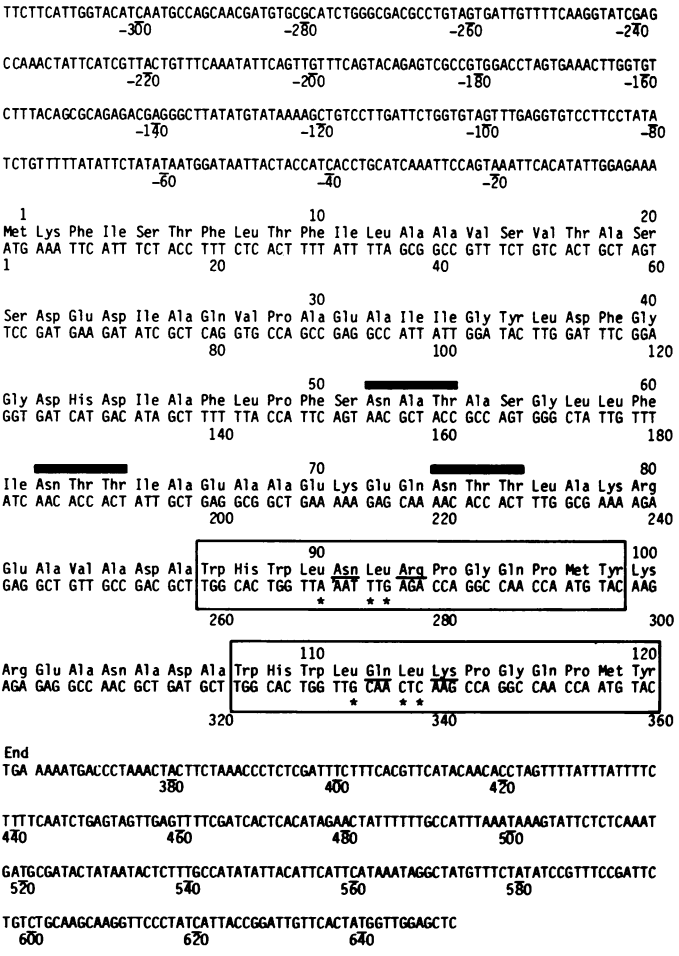

Figure 4. Nucleotide sequence of MFα2 gene and its nontranslated 5' and 3' flanking regions. The underlined amino acids indicate differences between the two pheromone copies encoded by the MFα2 gene. See Fig. 3 for other details.

residues between this hydrophobic sequence and the first repeat unit include 3 possible recognition sites for N-glycosylation (37). The organization of the pheromone gene contained in the p53 clone is identical to the MFα gene recently described by Kurjan and Herskowitz (15). This gene differs from MFα at 4 positions. It contains T (instead of C) residues at positions -8, -7, and 125 and an A (instead of C) residue at position 604. Because of the difference at position 125, there is a TTA (Leu) rather than TCA (Ser) codon

at amino acid position 42. We have designated the gene contained in p53 as MFα1.

A different α-factor gene, MFα2, is present in the p56 clone. The organization of this gene (Fig. 4) is similar, but not identical, to the MFα1. The α-factor encoded by this gene is apparently made as a precursor protein of 120 amino acid residues containing two copies of the pheromone. One of the α-pheromone tridecapeptides contained in the putative precursor is identical to the pheromone copies encoded by the MFα1 gene, whereas the second copy contains a Gln → Asn and a Lys → Arg substitution. As is the case with MFα1, this putative α-factor precursor also starts with a highly hydrophobic sequence (21 amino acids) which is followed by a sequence that contains 3 possible recognition sites for N-glycosylation (Fig. 4). However, although there are regions of homology (see Discussion) the primary sequence of the "pre-pro" portions of the two putative precursors are quite dissimilar.

Thus, the yeast α-factor is apparently synthesized in the pre-pro forms with structures similar to some of the known mammalian multivalent protein precursors (38-43).

Isolated α-Factor Genes are Functional. We used the property of growth inhibition of a cells by α-factor to test whether or not the pheromone genes contained in the cloned 1.7 kbp EcoRI and 1.8 kbp HindIII fragments are functional. We reasoned that if an active α-factor pheromone gene were present in a plasmid, significantly more pheromone would be synthesized in cells containing the multi-copy plasmid than in cells containing only the chromosomal copy (or copies) of the gene. The enhanced levels of the α-factor could then be detected by an increase in the area of nongrowth in a lawn of responsive a cells. The 1.7 kbp fragment, isolated from EcoRI-digested p53 DNA, and the 1.8 kbp fragment, isolated from HindIII-digested p56 DNA, were separately ligated to a pBR322-based vector plasmid which contained the yeast selectable marker TRP1 and the yeast origin of replication from the 2μm yeast plasmid (44). Yeast strain 20B-12 was separately transformed with these plasmids and with a control plasmid that lacked DNA sequences coding for the α-factor. The transformants were then compared for pheromone production. As seen in Fig. 4, the transformants containing MFα1 or MFα2 coding sequences on plasmids produce significantly more α-factor than the same strain transformed with the control plasmid. As expected, the strain of a mating type produces no α-factor. We conclude that the 1.7 kbp EcoRI (MFα1) and 1.8 kbp HindIII (MFα2) fragments contain

active α-factor pheromone genes. Our result with MFα1 is consistent with that of Kurjan and Herskowitz (15), as this gene corresponds to the gene described by them.

DISCUSSION

We have identified and isolated two genes coding for the α-factor of the yeast Saccharomyces cerevisiae. We exploited the fact that α-pheromone arrests the growth of a cells at G1 period of the cell cycle (2,4) to show that the isolated genes are functional (Fig. 5).

The primary structure of the proteins deduced from the nucleotide sequence of the α-factor gene shows that the yeast α-pheromone is processed from large precursor molecules. A comparison of the two putative α-factor precursors is shown in Fig. 6. The organization of these precursors is strikingly similar to that of certain mammalian precursors for neuroendocrine peptides. Thus, like the proopiomelanocortin (38,40), proenkephalin (41-43), and prodynorpnin (39), the yeast precursors contain multiple peptide units destined for secretion. In all these precursors the secreted unit is contained on the C-terminal half of the precursors. The N-terminal half of the molecules carry possible glycosylation sites. As is the case for the mammalian multifunctional precursors, glycosylation may be involved in the correct processing of the α-factor precursor. However, the actual processing steps for the yeast precursor seem to be unpredictably different from those of mammalian precursor proteins. Whereas the pairs of basic residues (Lys-Arg) providing sites for release directly flank the secreted peptide in the mammalian precursors, cleavage at these sites in the α-factor precursor would release the pheromone units with several additional amino acids at the N-terminus (see Figs. 3 and 4). These N-terminal extensions would consist of repeating -X-Ala- sequences in the precursors encoded by both MFα1 and MFα2 genes. Recent experiments of Thorner and his collaborators (14,45) indicate that the last step in the processing of the α-factor precursors is the removal of these sequences by dipeptidyl amino peptidases. The dipeptidyl amino peptidase A characterized by these authors is specific for -X-Ala- sequences (14). The bee venom melittin (46) and the frog skin caerulein (47) precursors are apparently processed by similar mechanisms.

We have determined the nucleotide sequence of the region coding for two putative α-factor precursor protein together with their 5' and 3' flanking regions. A detailed discussion of these flanking regions is premature since the transcription initiation and termination sites have not yet been

Table 1. Codon Usage in MFα Genes

		α1	α2			α1	α2			α1	α2			α1	α2
Phe	UUU	4	4	Ser	UCU	1	2	Tyr	UAU	0	0	Cys	UGU	0	0
Phe	UUC	2	3	Ser	UCC	3	1	Tyr	UAC	5	3	Cys	UGC	0	0
Leu	UUA	6	3	Ser	UCA	1	0	End	UAA	1	0	End	UGA	0	1
Leu	UUG	5	5	Ser	UCG	0	0	End	UAG	0	0	Trp	UGG	8	4
Leu	CUU	0	0	Pro	CCU	4	0	His	CAU	4	1	Arg	CGU	0	0
Leu	CUC	0	2	Pro	CCC	1	0	His	CAC	0	2	Arg	CGC	0	0
Leu	CUA	3	1	Pro	CCA	6	6	Gln	CAA	8	4	Arg	CGA	0	0
Leu	CUG	2	0	Pro	CCG	1	0	Gln	CAG	1	1	Arg	CGG	0	0
Ile	AUU	4	5	Thr	ACU	4	4	Asn	AAU	2	1	Ser	AGU	0	3
Ile	AUC	1	2	Thr	ACC	0	4	Asn	AAC	3	4	Ser	AGC	2	0
Ile	AUA	1	1	Thr	ACA	3	0	Lys	AAA	6	3	Arg	AGA	5	3
Met	AUG	5	3	Thr	ACG	1	0	Lys	AAG	3	2	Arg	AGG	0	0
Val	GUU	3	2	Ala	GCU	15	10	Asp	GAU	5	5	Gly	GGU	1	1
Val	GUC	2	1	Ala	GCC	4	6	Asp	GAC	2	2	Gly	GGC	4	2
Val	GUA	1	0	Ala	GCA	5	0	Glu	GAA	14	2	Gly	GGA	0	2
Val	GUG	0	1	Ala	GCG	0	3	Glu	GAG	1	5	Gly	GGG	3	1

determined. However, the following features of the sequence should be noted. There are A residues at position -3 in both MFα1 and MFα2 clones, as has been found for all yeast nuclear genes that are transcribed by RNA polymerase II (48). There are only two G residues among the 44 nucleotides immediately preceding ATG in MFα1 and the MFα2 gene contains five G residues among the 54 nucleotides preceding the initiation codon. Such a paucity of G residues in this region has also been seen in other yeast genes (48). As noted by Zalkin and Yanofsky (49), most yeast genes contain the sequence PuCACACA four to 15 nucleotides upstream from the translation initiation codon. A similar sequence (ATACACA) is present at positions -19 to -25 in the α-factor gene, MFα1. The analogous sequence in MFα2 gene would be TCACATA at positions -10 to -16. Both α-factor genes have a TATA box at identical positions. The canonical TATA sequence (TATAAAA) is present at positions -128 to -122 in the MFα2 gene (Fig. 4) and a similar sequence (TATATAA) is found at the same positions in the MFα1 gene (Fig. 3). An

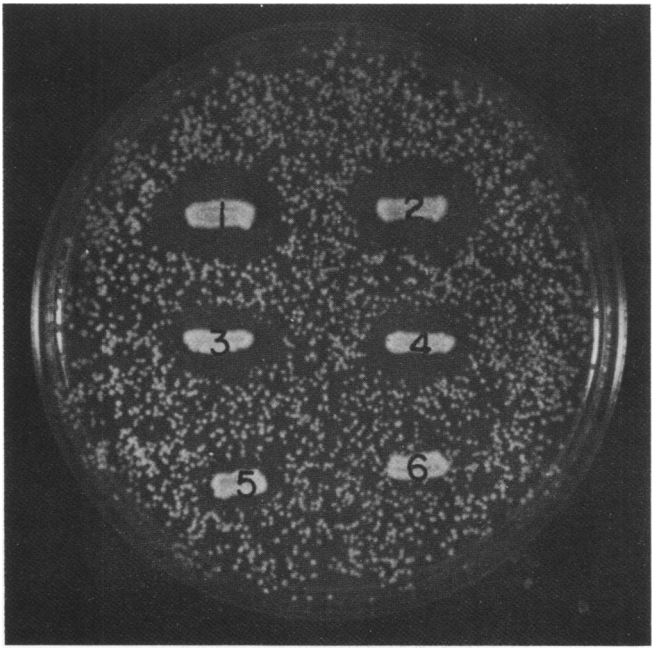

Figure 5. Functional test of the α-factor gene. Cells of six strains were streaked on a minimal agar plate containing 0.5 percent casamino acids. The plate had been previously seeded with approximately 10^5 cells of strain 483. The photograph was taken after three days of growth at 23°. Streaks: 1 and 2, two independent 20B-12 transformants with plasmid containing MFα1; 3 and 4, two independent 20B-12 transformants with plasmid containing MFα2; 5, 20B-12 transformant with control plasmid; 6, control strain of a mating type. The clear areas around the streaks on the lawn of growing 483 cells represent inhibition of a cells by secreted α-pheromone.

examination of the 5' and 3' flanking regions of the two genes shows no other striking homology between the two sequences.

It is of interest to examine the codon usage in the α-factor genes since the cloning strategy made use of codon usage frequencies from three yeast genes (Figure 1). As shown in Table 1, the codon selection for various amino acids in the α-factor genes is nonrandom. However, codon selection is not as strongly biased as in some other yeast genes (27). There are five cases in MFα1 and three cases in MFα2, marked by asterisks in Figs. 3 and 4, where different codons for the same amino acid are utilized in the α-factor coding regions. Similar differences in codon selection are also present in

```
                1                   10                  20
MFα1            MetArgPheProSerIlePheThrAlaValLeuPheAlaAlaSerSerAlaLeuAlaAla
MFα2            MetLysPheIleSerThrPheLeuThrPhe----------IleLeuAlaAla
Consensus       Met---Phe---Ser---Phe----------LeuAlaAla

                                30                  40
MFα1            ProValAsnThr---ThrThrGluAspGluThrAlaGlnIleProAlaGluAlaValIle
MFα2            ValSerValThrAlaSerSerAspGluAspIleAlaGlnValProAlaGluAlaIleIle
Consensus       ---------Thr----------AlaGln---ProAlaGluAla---Ile

                                50                  60
MFα1            GlyTyrLeuAspLeuGluGlyAspPheAspValAlaValLeuProPheSerAsnSerThr
MFα2            GlyTyrLeuAspPheGlyGlyAspHisAspIleAlaPheLeuProPheSerAsnAlaThr
Consensus       GlyTyrLeuAsp------GlyAsp---Asp---Ala---LeuProPheSerAsn---Thr

                                70                  80
MFα1            AsnAsnGlyLeuLeuPheIleAsnThrThrIleAlaSerIleAlaAlaLysGluGluGly
MFα2            AlaSerGlyLeuLeuPheIleAsnThrThrIleAlaGluAlaAlaGluLysGluGlnAsn
Consensus       ------GlyLeuLeuPheIleAsnThrThrIleAla------Ala---LysGlu------

                                90                  100
MFα1            ValSerLeuAspLysArgGluAlaGlu------AlaTrpHisTrpLeuGlnLeuLysPro
MFα2            ThrThrLeuAlaLysArgGluAlaValAlaAspAlaTrpHisTrpLeuAsnLeuArgPro
Consensus       ------Leu---LysArgGluAla---------AlaTrpHisTrpLeu---Leu---Pro

                                110                 120
MFα1            GlyGlnProMetTyrLysArgGluAlaGluAlaGluAlaTrpHisTrpLeuGlnLeuLys
MFα2            GlyGlnProMetTyrLysArgGluAlaAsnAlaAspAlaTrpHisTrpLeuGlnLeuLys
Consensus       GlyGlnProMetTyrLysArgGluAla---Ala---AlaTrpHisTrpLeuGlnLeuLys

                                130                 140
MFα1            ProGlyGlnProMetTyrLysArgGluAlaAspAlaGluAlaTrpHisTrpLeuGlnLeu
MFα2            ProGlyGlnProMetTyr
Consensus       ProGlyGlnProMetTyr----------------------------------------

                                150                 160
MFα1            LysProGlyGlnProMetTyrLysArgGluAlaAspAlaGluAlaTrpHisTrpLeuGln
Consensus       ------------------------------------------------------------

                            168
MFα1            LeuLysProGlyGlnProMetTyr
Consensus       -------------------------
```

Figure 6. Comparison of amino acid sequences of the putative α-factor precursors encoded by the MFα1 and MFα2 genes. Various gaps were created to align the sequences with maximum homology.

the regions that separate the coding blocks of the α-factor.

No α-factor with the amino acid sequence corresponding to the first copy in MFα2 precursor (see Fig. 4) has been reported, although our results (Fig. 5) indicate that MFα2 is a functional gene. This copy of the α-factor differs from others at two positions by substitutions of amino acids with similar properties (Gln → Asn and Lys → Arg). It is likely that this species of α-factor has not been detected because, assuming that the yeast genome contains only the two α-factor genes described here and that both are expressed equally, it would constitute a small portion (17 percent) of the pheromone molecules. Another possibility that should be considered is that

all the MFα2 activity seen here is attributable to the second copy of the α-factor (which is identical to the four copies present in MFα1) contained in the putative precursor. Still another possibility to be considered is that there is no expression of the chromosomal MFα2 gene. Further experiments with the cloned α-factor genes will be helpful in studying these possibilities as well as in elucidating the mechanisms involved in the processing of the pheromone precursor proteins and the regulation of their expression.

ACKNOWLEDGMENTS

We thank Joel S. Hayflick for his help in DNA sequence determination and Jeanne Arch for help in the preparation of the manuscript.

REFERENCES

1. Levi, J.D. (1956) Nature 177, 753-754.
2. Duntze, W., MacKay, V.L. and Manney, T.R. (1970) Science 168, 1472-1473.
3. Wilkinson, L.E. and Pringle, J.R. (1974) Exp. Cell Res. 89, 175-187.
4. Bücking-Throm, E., Duntze, W., Hartwell, L.H. and Manney, T.R. (1973) Exp. Cell Res. 76, 99-110.
5. Hartwell, L.H. (1973) Exp. Cell Res. 76, 111-117.
6. MacKay, V.L. and Manney, T.R. (1974a) Genetics 76, 255-271.
7. MacKay, V.L. and Manney, T.R. (1974b) Genetics 76, 273-288.
8. Stotzler, D. and Duntze, W. (1976) Eur. J. Biochem. 65, 257-262.
9. Stotzler, D., Kiltz, H.H. and Duntze, W. (1976) Eur. J. Biochem. 69, 397-400.
10. Sakurai, A., Tamura, N., Yanagishima, N. and Shimoda, C. (1976) Agric. Biol. Chem. 40, 1057-1058.
11. Ciejek, E., Thorner, J. and Geier, M. (1977) Biochem. Biophys. Res. Commun. 78, 952-961.
12. Tanaka, T., Kita, H., Murakami, T. and Narita, K. (1977) J. Biochem. 82, 1681-1687.
13. Scherer, G., Haag, G. and Duntze, W. (1974) J. Bacteriol. 119, 386-393.
14. Julius, D., Blair, L., Brake, A., Sprague, G. and Thorner, J. (1983) Cell 32, 839-852.
15. Kurjan, J. and Herskowitz, I. (1982) Cell 30, 933-943.
16. Backman, K., Ptashne, M. and Gilbert, W. (1976) Proc. Natl. Acad. Sci. USA 73, 4174-4178.
17. Sherman, F., Fink, G. and Lawrence, C. (1979) Methods in Yeast Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. Miller, J.H. (1972) Experiments in Molecular Genetics, p. 433, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. Mandel, M. and Higa, A. (1970) J. Mol. Biol. 53, 159-162.
20. Hinnen, A., Hicks, J.B. and Fink, G.R. (1978) Proc. Natl. Acad. Sci. USA 75, 1929-1933.
21. Beggs, J.D. (1978) Nature 275, 104-109.
22. Clewell, D.B. and Helinski, D.R. (1970) Biochemistry 9, 4428-4440.
23. Birnboim, H.C. and Doly, J. (1979) Nucleic Acid Res. 7, 1513-1523.
24. Crea, R. and Horn, T. (1980) Nucleic Acid Res. 8, 2331-2348.
25. Holland, J.P. and Holland, M.J. (1979) J. Biol. Chem. 254, 9839-9845.

26. Holland, J.P. and Holland, M.J. (1980) J. Biol. Chem. 255, 2596-2605.
27. Bennetzen, J.L. and Hall, B.D. (1982) J. Biol. Chem. 257, 3026-3031.
28. Struhl, K., Stinchcomb, D.T., Scherer, S. and Davis, R.W. (1979) Proc. Natl. Acad. Sci. USA 76, 1035-1039.
29. Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3965.
30. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
31. Smith, A.J.H. (1980) Methods Enzymol. 65, 499-560.
32. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
33. Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucleic Acid Res. 9, 309-321.
34. Sanger, F. and Coulson, A.R. (1980) FEBS Letters 87, 107-110.
35. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
36. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. (1977) Gene 2, 95-119.
37. Struck, D.K., Lennarz, W.J. and Brew, K. (1978) J. Biol. Chem. 253, 5786-5794.
38. Mains, R.E., Eipper, B.A. and Ling, N. (1977) Proc. Natl. Acad. Sci. USA 74, 4826-4830.
39. Kakidani, H., Furutani, Y., Takahashi, H., Noda, M., Morimoto, Y., Hirose, T., Asai, M., Inayama, S., Nakanishi, S. and Numa, S. (1982) Nature 298, 245-249.
40. Nakanishi, S., Inoue, A., Kita, T., Nakamura, N., Chang, A.C.Y., Cohen, S.N. and Numa, S. (1979) Nature 278, 423-427.
41. Gubler, U., Seeburg, P., Hoffman, B.J., Gage, L.P. and Udenfriend, S. (1982) Nature 295, 206-208.
42. Comb, M., Seeburg, P.H., Adelman, J., Eiden, L. and Herbert, E. (1982) Nature 295, 663-666.
43. Noda, M., Furutani, Y., Takahashi, H., Toyosato, M., Hirose, T., Inayama, S., Nakanishi, S. and Numa, S. (1982) Nature 295, 202-206.
44. Broach, J.R., Strathern, J.N. and Hicks, J.B. (1979) Gene 10, 157-166.
45. Barnes, D., Blair, L., Brake, A., Church, M., Julius, D., Kunisawa, R., Lotko, J., Stetler, G. and Thorner, J. (1982) in Berkeley Workshop on Recent Advances in Yeast Molecular Biology: Recombinant DNA, Esposito, M.S., Mortimer, R.K., Bruschi, C., Schild, D. and Herskowitz, I., Eds., pp. 295-305, Lawrence Radiation Laboratory, University of California, Berkeley, California.
46. Kreil, G., Haiml, L. and Suchanek, G. (1980) Eur. J. Biochem. 111, 49-58.
47. Hoffmann, W., Bach, T.C., Seliger, H. and Kreil, G. (1983) EMBO J. 2, 111-114.
48. Ammerer, G., Hitzeman, R., Hagie, F., Barta, A. and Hall, B.D. (1981) in Recombinant DNA, Proceedings of the Third Cleveland Symposium on Macromolecules, Walton, A.G., Ed., pp. 185-197, Elsevier Scientific Publishing Co., Amsterdam, Netherlands.
49. Zalkin, H. and Yanofsky, C. (1982) J. Biol. Chem. 257, 1491-1500.