MOLECULAR AND CELLULAR BIOLOGY, Sept. 1996, p. 4700–4709 Vol. 16, No. 9
0270-7306/96/$04.00+0
Copyright © 1996, American Society for Microbiology

# Yeast α Mating Factor Structure-Activity Relationship Derived from Genetically Selected Peptide Agonists and Antagonists of Ste2p

JOHN P. MANFREDI,[1]* CHRISTINE KLEIN,[1] JUAN J. HERRERO,[1] DEVON R. BYRD,[2] JOSHUA TRUEHEART,[1] WILLIAM T. WIESLER,[1] DANA M. FOWLKES,[1]† AND JAMES R. BROACH[1,3]

*Cadus Pharmaceutical Corporation, Tarrytown, New York 10591-6705[1]; Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599[2]; and Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544[3]*

Received 25 March 1996/Returned for modification 3 May 1996/Accepted 14 June 1996

**α-Factor, a 13-amino-acid pheromone secreted by haploid α cells of *Saccharomyces cerevisiae*, binds to Ste2p, a seven-transmembrane, G-protein-coupled receptor present on haploid a cells, to activate a signal transduction pathway required for conjugation and mating. To determine the structural requirements for α-factor activity, we developed a genetic screen to identify from random and semirandom libraries novel peptides that function as agonists or antagonists of Ste2p. The selection scheme was based on autocrine strains constructed to secrete random peptides and respond by growth to those that were either agonists or antagonists of Ste2p. Analysis of a number of peptides obtained by this selection procedure indicates that Trp1, Trp3, Pro8, and Gly9 are important for agonist activity specifically. His2, Leu4, Leu6, Pro10, a hydrophobic residue 12, and an aromatic residue 13 are important for both agonist and antagonist activity. Our results also show that activation of Ste2p can be achieved with novel, unanticipated combinations of amino acids. Finally, the results suggest the utility of this selection scheme for identifying novel ligands for mammalian G-protein-coupled receptors heterologously expressed in *S. cerevisiae*.**

Haploid *Saccharomyces* cells of mating type α secrete a 13-amino-acid pheromone, α-factor (Trp1-His2-Trp3-Leu4-Gln5-Leu6-Lys7-Pro8-Gly9-Gln10-Pro11-Met12-Tyr13), that can bind and stimulate Ste2p, a G-protein-coupled receptor expressed on the surface of haploid cells of mating type **a** (reviewed in reference 26). Binding of α-factor to Ste2p activates the pheromone signaling pathway by promoting on the cytoplasmic side of the plasma membrane dissociation of a coupled heterotrimeric G protein into its constituent α subunit and βγ complex. The βγ complex in conjunction with Cdc42p initiates a kinase cascade that results in activation of mitogen-activated protein kinase homologs, Kss1p and Fus3p (reviewed in reference 17). These activated kinases promote both cell cycle arrest, through inactivation of the cyclin B/p34$^{cdc28}$ complex in a process mediated by Far1p, and transcription of a variety of genes required for mating and conjugation, mediated by the Ste12p transcriptional activator. Activation of this signal transduction pathway is essential for productive mating, and mutations that attenuate this pathway result in sterility.

Current knowledge of α-factor suggests a complex structure-function relationship. Although determinants of its various biological activities are not restricted to discrete peptide regions, the determinants may be more heavily concentrated in certain domains. For example, residues that initiate signaling may be concentrated in the N terminus, while those that mediate binding may dominate the C-terminal region. This suggestion is based on observations that removal or substitution of Trp1, His2, or Trp3 results in loss of biological activity that exceeds the relatively small reductions in binding affinity (10, 22). For example, removal of Trp1 to form des-Trp1-α-factor yields a peptide with about 10% the affinity but less than 1% the bioactivity of α-factor. Further mutation of this truncated peptide to des-Trp1-[Phe3]-α-factor or des-Trp1-[Ala3]-α-factor causes complete loss of bioactivity with no further diminution of affinity (22). A relatively minor role for the N-terminal region in binding compared with agonist efficacy is further indicated by the observation that removal of both Trp1 and His2 produces an analog that shows about 25% the binding affinity but less than 0.1% the bioactivity of α-factor (10, 22). In contrast, coincident removal of the C-terminal Met12 and Tyr13 leads to complete loss of agonist activity that is accounted for by loss of affinity for Ste2p (10). While such data support a model of α-factor in which the N and C termini are largely devoted to agonist efficacy and binding, respectively, additional data indicate that such a model is a crude approximation. Thus, it has been observed that deletion of Tyr13 alone yields a peptide that acts synergistically with α-factor (10). Also, the middle region of α-factor is critical for biological activity, perhaps as a result of the formation of a transient type II β turn (12, 14, 30).

We have further explored the structure-function relationship of α-factor by using a genetic system to select α-factor analogs that act as either agonists or antagonists of Ste2p. Our results identify novel peptide agonists and antagonists of Ste2p whose sequences provide information for the developing model of α-factor structure and activity. The genetic selection strategy may be used to discover novel ligands for any G-protein-coupled receptor functionally expressed in *Saccharomyces cerevisiae.*

## MATERIALS AND METHODS

**Strains.** *S. cerevisiae* CY1455 (*MAT***a** *bar1*::*hisG far1-1 fus1-HIS3 ste14*::*TRP1 ura3 trp1 leu2 his3*) and CY1892 (*MAT***a** *bar1*::*hisG fus1-HIS3 ura3 trp1 leu2 his3 SUC*$^+$) were obtained by standard genetic techniques, with IH2512 (provided by I. Herskowitz), SY1390 (27) (provided by G. Sprague), SM1188 (24) (provided by S. Michaelis), and JY512 (provided by J. Trueheart) serving as the original sources of the *far1-1*, *fus1-HIS3*, *ste14*, and *bar1* alleles, respectively. CY2797 (*MAT***a** *tbt1-1 bar1*::*hisG far1-1 fus1-HIS3 ste14*::*TRP1 ste2Δ1154 ura3 trp1 leu2*

* Corresponding author. Mailing address: Cadus Pharmaceutical Corporation, 777 Old Saw Mill River Rd., Tarrytown, NY 10591-6705. Phone: (914) 345-3344. Fax: (914) 345-3565. Electronic mail address: 71053.3661@compuserve.com.

† Present address: Department of Pathology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.

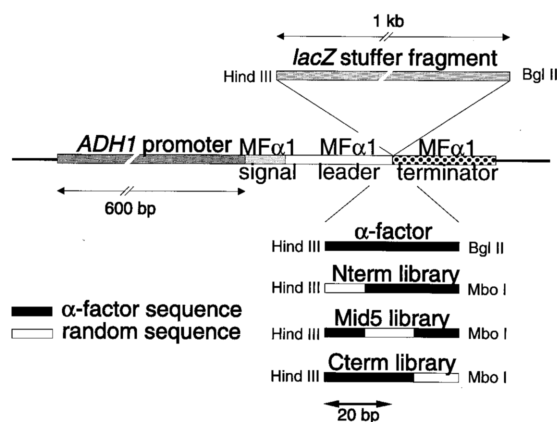


FIG. 1. Design of cassette for expression of α-factor and the semirandom α-factor libraries. Expression of α-factor or the semirandom libraries in the context of the α-factor prepro sequence is accomplished by cloning the peptide-encoding oligonucleotides, containing *Hin*dIII and either *Bcl*I or *Mbo*I ends, into *Hin*dIII- and *Bgl*II-digested vector, thereby replacing a *lacZ*-containing fragment in the vector with the oligonucleotides. The *Bcl*I-*Bgl*II and *Mbo*I-*Bgl*II junctions encode an in-frame stop codon following the 13th codon of each oligonucleotide. Processing by Kex2p and Ste13p of the primary translation products encoded by the resulting plasmids is predicted to generate tridecapeptides identical to α-factor over the entire length (α-factor), the C-terminal nine amino acids (Nterm library), the N- and C-terminal four amino acids (Mid5 library), or the N-terminal nine amino acids (Cterm library).

*his3*) was obtained from CY1455 by the two-step disruption (4) of *STE2* with plasmid CP1154, constructed as follows. A 4.3-kb *Bam*HI fragment containing *STE2* was isolated from YEp24-STE2 (provided by J. Thorner) and subcloned into pALTER (Promega) for insertion of *Spe*I sites 7 nucleotides upstream of the start codon and 9 nucleotides downstream of the stop codon. The mutagenized *Bam*HI fragment was subcloned into the *Bam*HI site of YIp19, and the resulting plasmid was cut with *Spe*I and religated to yield CP1154.

**Media.** Media were prepared essentially as described previously (25). SC-Ura and SC-His are yeast nitrogen base containing 2% glucose and an amino acid and adenine supplement (0.5 g/liter) lacking uracil and histidine, respectively; SC-UraHis lacks histidine as well as uracil. 3(1,2,4)-Aminotriazole (Sigma, St. Louis, Mo.) was added as indicated to titrate the activity of the HIS3 reporter enzyme (15).

**Vector construction.** The vector used for peptide expression by yeast strains was constructed as follows. A fragment comprising approximately 600 bp of the upstream activation sequence of the *ADH1* promoter (−10 to about −600 relative to the translational start) was obtained from pAAH5 (1) and cloned as an *Sph*I-to-*Hin*dIII fragment into the *Bam*HI and *Hin*dIII sites in the polylinker of pRS426 (8). A fragment that contains 5 bp of the *MFα1* mRNA leader followed by sequence encoding a modified prepro segment of the α-factor precursor was then cloned into the *Hin*dIII site such that expression of sequences fused to the MFα1 prepro leader is directed by the *ADH1* promoter (Fig. 1). The prepro-encoding segment is modified as a result of mutation of TA (nucleotides 247 and 248 relative to the start of the *MFα1* coding sequence) to CT, thereby introducing an *Afl*II site for possible insertion of oligonucleotides while mutating Asp83 to Leu. Processing of the α-factor precursor is tolerant to mutations at this residue (7). The prepro sequences terminate in a *Hin*dIII site, as they do in native *MFα1* and *MFα2*, followed by a 1-kb stuffer fragment encoding the α fragment of *lacZ* joined via a *Bgl*II site at its 3′ end to approximately 300 bases of the *MFα1* 3′ untranslated sequence. The stuffer fragment facilitates library construction by allowing more complete isolation of *Hin*dIII- and *Bgl*II-cut vector from singly cut vector. At the same time, vector background in the libraries can be determined from the relative numbers of blue and white bacterial colonies following transformation with the plasmid libraries.

**Construction of libraries.** Three chemically synthesized, partially random oligonucleotide libraries were constructed to determine structural features of the N-terminal, middle, and C-terminal regions of α-factor that contribute to its bioactivity. The N-terminal (Nterm) library was made with oligonucleotides that use triplets NNK to encode residues 2, 3, and 4 of the mature peptide (N is any nucleotide and K is either G or T), triplets TNK to encode the first residue while preserving the *Hin*dIII site, and codons 5 through 13 of genuine α-factor. Thus, this library encodes amino acids C, F, L, S, W, and Y at residue 1 of the peptide, any amino acid at residues 2, 3, and 4, and the normal α-factor residues at the remaining nine positions (viz., QLKPGQPMY). The double-stranded oligonucleotides containing the random nucleotides were constructed as follows. First, 1 nmol each of the two single-stranded oligonucleotides 5′-GCCGTCAGTAAAGCTTNKNNKNNKNNKCAGCTAAAACCTGGCCAGCCTATGTACTGATCAGTCTGTGACGC-3′ and 5′-GCGTCACAGACTGATCA-3′ (N is an equimolar mixture of A, C, G, and T, and K is equimolar G and T) were annealed by combining in 200 μl of 40 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 0.1 mg of bovine serum albumin per ml, and 0.5 mM dithiothreitol, heating the mixture to 65°C for 20 min, and then slowly cooling it to 37°C. The partially double-stranded oligonucleotides were then filled in by adding 5 μl of 10 mM deoxynucleoside triphosphates and 3 μl of Sequenase (13 U/μl; United States Biochemical Corp.) followed by successive incubations at 37°C for 30 min and 65°C for 20 min. The DNA was then cut with *Hin*dIII and *Mbo*I, the digested DNA was gel isolated, and approximately 5 pmol was ligated to 1.5 pmol of *Hin*dIII- and *Bgl*II-cut expression vector (Fig. 1) in 1 ml at 16°C overnight. The *Mbo*I-*Bgl*II junction creates a stop codon immediately after the triplet encoding residue 13. Transformation of *Escherichia coli* JS5 with the ligation mixture yielded about $10^8$ ampicillin-resistant clones that should contain practically all of the $6 \times 10^5$ possible nucleotide sequences. Three of the possible nucleotide sequences should encode authentic α-factor, which itself is one of about $6.4 \times 10^4$ possible peptides encoded by the library. The quality of the library was evaluated by sequencing the peptide-encoding inserts in plasmids recovered from 20 randomly chosen yeast clones that had been transformed with the *E. coli*-amplified library. All 20 inserts encode tridecapeptides, and the frequencies of amino acids at the randomized positions were not significantly different from those predicted by chance alone.

The Mid5 library was made with oligonucleotides that use triplets NNK in the region encoding amino acids 5 through 9; other triplets encode α-factor amino acids. The double-stranded oligonucleotides were constructed as described for the Nterm library, but with the following single-stranded oligonucleotides: 5′-GCCGTCAGTAAAGCTTGGCATTGGTTGNNKNNKNNKNNKNNKCAGCCTATGTACTGATCAGTCTGTGACGC-3′ and 5′-GCGTCACAGACTGATCA-3′. As with the Nterm library, the oligonucleotides used to construct the Mid5 library were digested with *Hin*dIII and *Mbo*I and cloned into the *Hin*dIII- and *Bgl*II-cut expression vector (Fig. 1). The theoretical complexity of this library is $3.4 \times 10^7$ different nucleotide sequences encoding $3.2 \times 10^6$ different peptides. Transformation of *E. coli* JS5 yielded about $10^7$ ampicillin-resistant clones, which should contain about 25% of the possible nucleotide sequences, encoding about 95% of the $3.2 \times 10^6$ possible tridecapeptides. All 20 randomly chosen yeast colonies transformed with the *E. coli*-amplified library were found to carry plasmids that encode tridecapeptides of the expected semirandom sequences, and the frequencies of amino acids at the randomized positions are not significantly different from that predicted by chance alone.

The Cterm library encodes α-factor amino acids at residues 1 through 9 and random, NNK-encoded amino acids at residues 10 through 13. As with the Nterm and Mid5 libraries, the double-stranded oligonucleotides constructed from two single-stranded oligonucleotides (in this case 5′-GCCGTCAGTAAAGCTTGGCATTGGTTGCAGCTAAAACCTGGCNNKNNKNNKNNKTGATCAGTCTGTGACGC-3′ and 5′-GCGTCACAGACTGATCA-3′) were digested with *Hin*dIII and *Mbo*I and cloned into *Hin*dIII- and *Bgl*II-cut expression vector (Fig. 1). Transformation of *E. coli* JS5 yielded about $4 \times 10^8$ ampicillin-resistant clones. Sequencing of plasmids isolated from randomly chosen yeast transformants, however, indicated that the quality of the library was low. Specifically, none of 30 plasmids isolated from unselected yeast encode tridecapeptides with the expected semirandom structure. Rather, most of these plasmids carry frameshift mutations within the first four codons of the inserts; others show substitution of several N-terminal codons; and one contains a 6-bp substitution of the expected Leu4 codon. As a result, the diversity of this library is greater than intended, and the chance that any particular peptide-encoding sequence is among the yeast transformants is lower than intended.

**Transformation protocols.** Following ligation of double-stranded oligonucleotides and the vector, the 1-ml ligation reaction mixture was extracted with phenol and chloroform, ethanol precipitated, and resuspended in 100 μl of water. Fifty microliters of the ligated DNA was then added to 200 μl of electrocompetent *E. coli* JS5, and 60 μl was aliquoted to four cuvettes for electroporation at 1.8 kV, 25 μF, and 200 Ω. The bacteria were then combined in 5 ml of LB supplemented with 20 mM glucose and incubated at 37°C for 1 h, and aliquots were removed to estimate library size; the remaining bacteria were added to 500 ml of LB with ampicillin for overnight growth at 37°C for isolation of library DNA. Yeast cells were transformed with the library DNA by electroporation as described by Becker and Guarente (2), using 1 to 25 μg of library DNA and approximately $10^8$ cells per 100 μl of transformation mixture.

**Chemically synthesized peptides.** Peptides were synthesized and prepared to 50 to 70% purity by Macromolecular Resources (Colorado State University, Fort Collins). Peptides were dissolved in dimethyl sulfoxide to 10 mg/ml, and dilutions were made in water to 0.5 mg/ml.

**Secondary assays for agonist and antagonist activities. (i) Liquid growth assay.** Potential autocrine strains, grown overnight in SC-Ura, were washed in water and resuspended in SC-His to an optical density at 600 of 0.012 (ca. 120,000 cells per ml). Samples of 90 μl were aliquoted to the wells of a 96-well plate that contain 10 μl of the same medium but with sufficient aminotriazole to give a final concentration 0.2 mM. The optical density at 600 nm was measured after 15 h at 30°C and compared with that of a strain of CY1455 that had been transformed with empty peptide vector, CP1625 (negative control), and with an α-factor-expressing plasmid, CP1219 (positive control). The reading of the negative control was subtracted from those of the experimental strain and positive

control, and the growth of the experimental strain is expressed as the ratio of its corrected growth to that of the positive control. Only those strains that exhibit positive growth ratios are considered to express Ste2p agonists.

**(ii) Paracrine arrest assay.** A suspension of strain CY1455 carrying agonist-encoding plasmids (ca. 10 μl containing $10^7$ cells) was spotted onto a YPD plate containing a lawn of strain CY1892, which undergoes growth arrest upon stimulation of Ste2p. In this halo assay, α-factor-like activity of peptides released by the spotted cells is observed as a clear zone of growth-arrested CY1892 surrounding the spotted cells.

**(iii) Paracrine activation assay.** A suspension of strain CY1455 carrying agonist-encoding plasmids (ca. 10 μl containing $10^7$ cells) was spotted onto an SC-His plate containing 0.2 mM aminotriazole and prespread with a lawn of strain CY1455, which grows on this medium upon stimulation of Ste2p. α-Factor-like activity of peptides released by the spotted cells is observed as a corona of growth surrounding the spotted cells.

**Statistical analyses.** We make the assumption that if there is no selection for a particular residue at a given position of a recovered peptide, the frequency of appearance of that residue at that position will be described by the binomial distribution. Accordingly, if genetic selection is operative, the frequency of appearance of the residue at that position will differ from the frequency predicted by the binomial distribution. This assumption allows standard statistical analyses to be used to determine if the observed frequency may be due to chance alone. We attempted to assess the reasonableness of this assumption by (i) sequencing plasmids derived from $Ura^+$ yeast clones that had been transformed with the library but not selected for growth on selective medium (i.e., medium containing α-factor or lacking histidine); (ii) tabulating the observed frequencies of encoded amino acids; and (iii) comparing these observed frequencies with the frequencies predicted by the oligonucleotide synthesis strategy used to construct the libraries. In this case, where no genetic selection is operative, the observed frequencies were not different from that predicted by the binomial distribution to be due to chance alone. This test of the appropriateness of the binomial distribution for our purposes is not particularly discriminating, and so the distribution must be regarded as an approximate description of random events in our system. But results that deviate significantly from predictions of the distribution nonetheless indicate the existence of nonrandom influences. We assume that these deviations result from the genetic selections that we impose on the system. Specifically, we base the predictions of the binomial distribution on the oligonucleotide synthesis strategy. We infer the existence of selection for or against a particular amino acid or class of amino acids at a particular position if the probability that the observed frequency can be described by the binomial distribution is less than $10^{-4}$. Such an apparently conservative requirement for significance is appropriate given the large number of comparisons made (typically 20 at each of the 13 positions).

For purposes of data analysis, the amino acids A, C, F, G, H, I, K, L, M, T, V, W, and Y are hydrophobic, and F, H, W, and Y are aromatic (29).

## RESULTS

**Identification of Ste2p agonists.** We have developed a protocol for identifying possible peptide agonists of the Ste2p receptor. For the discovery of Ste2p agonists, we used strain CY1455, which is engineered to grow under conditions of histidine starvation only if it synthesizes and secretes a Ste2p agonist. CY1455 contains *HIS3* coding sequences under the control of the genomic *FUS1* promoter, so that His3p expression is induced upon activation of the pheromone signaling pathway. In addition, this strain carries a *far1* mutation to prevent growth arrest consequent to activation of the pheromone response pathway, a *bar1* mutation to reduce degradation of α-factor and other similar peptides, and a *ste14* mutation to reduce background signaling through the pheromone response pathway.

The protocol for identification of Ste2p agonists is schematized in Fig. 2. CY1455 ($2 \times 10^8$ cells) was transformed with 1 μg of each of the plasmid libraries described in Materials and Methods and plated onto uracil-deficient solid medium to select for transformants. Following 3 days of growth at 30°C, the transformants were replica plated en masse to histidine-deficient medium containing 0.2 or 0.5 mM aminotriazole, a competitive inhibitor of the *HIS3* gene product (16, 28). These concentrations of aminotriazole are sufficient to eliminate the growth of unstimulated CY1455 that results from basal activity of the *FUS1* promoter. Transformants secreting peptides with agonist activity toward Ste2p were identified as growing colonies, due to autocrine activation of cells expressing α-factor-like peptides, surrounded by a corona of growth, due to paracrine activation of cells that do not themselves encode agonist peptides. An example of this pattern is shown in Fig. 3. Autocrine cells were colony purified by streaking onto uracil-deficient medium and retested for histidine prototrophy in the presence of 0.2 mM aminotriazole. Plasmids were isolated from the autocrine cells, amplified in *E. coli*, and reintroduced into the CY1455 test strain. The resulting transformants were tested for histidine prototrophy, and plasmids that enabled growth on histidine-deficient solid medium were introduced into CY2797, a derivative of CY1455 in which *STE2* was disrupted, to confirm dependence on Ste2p.

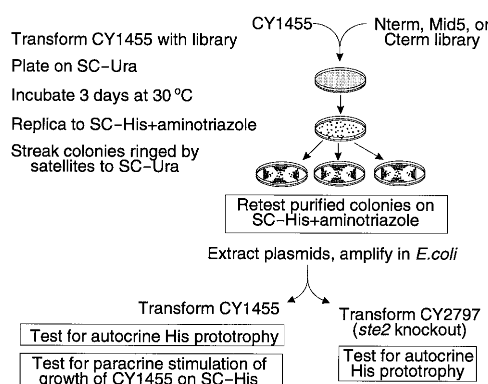


FIG. 2. Protocol for genetic selection of Ste2p agonists. Tests that are boxed in the diagram represent criteria that must be satisfied by candidate agonists.

Strains of CY1455 bearing plasmids that exhibited Ste2p agonist activity were further tested in three additional assays. In the first assay (liquid growth assay), suspensions of the strains were tested for growth in liquid media lacking histidine and containing aminotriazole concentrations of 0 and 0.2 mM. In the second assay (paracrine arrest assay), the Ste2p agonist activities of the plasmid-encoded peptides were tested for ability to arrest the growth of a supersensitive strain; specifically, agonist-secreting cells were spotted onto an overlay of tester cells that undergo growth arrest in the presence of α-factor. In the third assay (paracrine activation assay), agonist-secreting cells were spotted onto histidine-deficient medium on which a background lawn of CY1455 had been spread; peptides that are released by the agonist-secreting cells can stimulate the

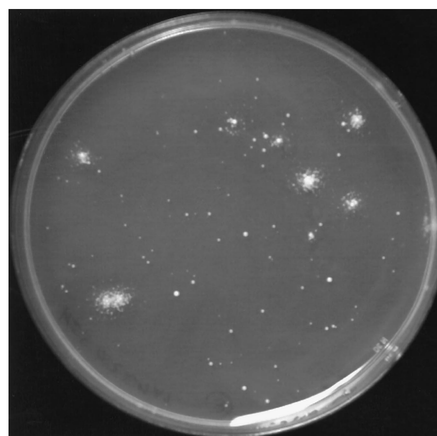

FIG. 3. Autocrine selection of peptide agonists of the Ste2p receptor. Clones secreting α-factor-like peptides were distinguished from constitutively activated mutant clones by a corona of growing cells that show paracrine response to secreted peptides.

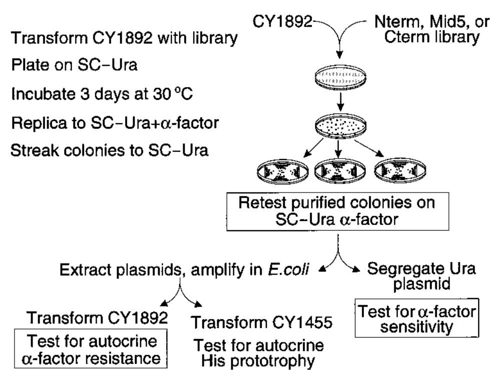


FIG. 4. Protocol for genetic selection of Ste2p antagonists. Tests that are boxed in the diagram represent criteria that must be satisfied by candidate antagonists.

background lawn, resulting in a corona of satellite colonies surrounding the spotted cells. Note that the first assay is an autocrine assay, while the second and third are paracrine assays. These paracrine assays reflect not only the agonist activity of a secreted peptide but also factors such as the permeability of the cell wall to the peptide and the diffusivity of the peptide in the solid agar-based media. The results of these assays are presented below.

**Identification of Ste2p antagonists.** The protocol for isolation of Ste2p antagonists is outlined in Fig. 4. Strain CY1892 carries the α-factor receptor coupled to the normal growth inhibition response, so that cells plated on medium containing 25 nM α-factor fail to grow. We anticipated that any cell producing a peptide that functioned as a Ste2p antagonist would interdict the growth arrest signal elicited by α-factor and allow the cell to grow. Accordingly, CY1892 ($3.5 \times 10^8$ cells) was transformed with 25 μg of the libraries described in Materials and Methods. Transformants were selected on medium lacking uracil and replica plated to medium lacking uracil but containing 25 nM α-factor. α-Factor-resistant colonies were isolated, colony purified, and assessed for the plasmid dependence of the α-factor resistance. Specifically, α-factor-resistant transformants were cured of episome and retested for sensitivity to α-factor. In addition, plasmids from each possible antagonist-producing strain were amplified in *E. coli* and reintroduced into naive CY1892, and the transformants tested for resistance to growth arrest by α-factor. Finally, the agonist activity of the putative antagonist was determined by transforming CY1455 with plasmids encoding candidate antagonists and measuring the growth of the peptide-secreting strain on histidine-deficient medium containing 0.2 mM aminotriazole. Antagonist-encoding plasmids that fail to confer autocrine histidine prototrophy are referred to as pure antagonists; those that confer detectable growth on histidine-deficient media are designated partial agonists.

**Ste2p agonists and antagonists isolated from the Nterm library.** Transformation of strain CY1455 with the Nterm library yielded $9 \times 10^5$ Ura$^+$ transformants, so that each peptide, including α-factor, should have been represented in yeast 14 times on average. Transformation of strain CY1892 gave $2 \times 10^6$ Ura$^+$ transformants, leading to the expectation that each peptide would have been represented about 30 times among these transformants. However, multiple isolations of the same peptide from the Nterm library were rare, even when all of the Ura$^+$ transformants were examined, suggesting that the reliability of the selection procedure in detecting the peptide-dependent phenotype is less than 100% and/or that plasmids encoding particular peptides are unevenly represented. Indeed, reconstruction experiments indicated that only 20% of cells containing an α-factor-encoding plasmid were recovered by the genetic screen for Ste2p agonists when present at a frequency of 1 α-factor-secreting cell in $10^5$. We expect that cells expressing peptides with weaker phenotypes would have been recovered at even lower frequencies, reflecting a bias of the selection system toward recovery of more potent agonists.

The Nterm library can encode amino acids C, F, L, S, W, and Y at residue 1 of the peptide, any amino acid at residues 2, 3, and 4, and the normal α-factor residues at the remaining 9 positions (viz., QLKPGQPMY). Genetic selection for Ste2p agonists identified 11 plasmids that confer Ste2p-dependent, autocrine histidine prototrophy, indicating that they encode peptides with α-factor-like activity (Table 1). Six of the 11 plasmids encode genuine α-factor, while the other five plasmids encode four novel α-factor analogs. The six plasmids encoding α-factor are presumably independent isolates, since the three possible α-factor-coding sequences are each represented one, two, and three times. Considering the peptides encoded by the 11 isolates, statistical analysis indicates significant bias for tryptophan, histidine, tryptophan, and leucine at positions 1, 2, 3, and 4, respectively.

Histidine prototrophy of the agonist-secreting strains was tested in histidine-deficient liquid culture containing 0 and 0.2 mM aminotriazole. (We have observed that an aminotriazole concentration of 0.2 mM in liquid media is roughly equivalent to 1 mM in solid media.) All strains exhibit greater growth than a strain carrying the empty expression vector, confirming that the strains synthesize peptides with α-factor-like activity (Table 1). However, in the absence of knowledge of the peptide concentrations to which Ste2p is exposed, these data provide no indication of the relative potencies of the secreted peptides. Also listed in Table 1 is a paracrine measure of agonist activity: the width of the zone of growth arrest of a lawn of strain CY1892 on which $10^7$ agonist-secreting cells were spotted (Fig. 5). Note that strains that exhibited equivalent growth rates in liquid media can give very different paracrine measures of agonist activity. Such apparent discrepancies between autocrine and paracrine assays of agonist activity may reflect differences in diffusion of the peptides in solid agar or in their cell wall permeabilities. The discrepancies may also reflect different thresholds for an agonist effect in the two assays.

Genetic selection for antagonists using the same Nterm library yielded many more positive clones than selection for agonists. From approximately 140 Ura$^+$ transformants of strain CY1892 that exhibited α-factor resistance, 38 were chosen for detailed analysis. Twenty-three of these transformants were shown to exhibit reproducible α-factor resistance that was plasmid dependent. These yielded 18 unique peptides encoded by 20 different plasmids, indicating that the frequency of detection of antagonist-secreting clones was less than 100%. Of the 20 plasmids with unique nucleotide sequences, 10 encoded peptides that display some agonist activity (see Table 3), as measured by aminotriazole-resistant growth on solid media. That is, 10 different plasmids were identified in the selection for antagonists that confer histidine prototrophy when transformed into strain CY1455. The remaining 10 unique plasmids, encoding nine different peptides, fail to confer histidine prototrophy on strain CY1455 and are therefore considered to encode pure antagonists (Table 2).

The predicted sequences of the pure antagonists collectively exhibit a strong preference for histidine in position 2 and for leucine in position 4. Similar preferences were observed for agonists isolated from this same library. No significant bias is

TABLE 1. Genetically selected Ste2p agonists

| Library | Predicted amino acid sequence[a] | Autocrine growth activity[b,c] | Paracrine growth arrest activity[c,d] |
|---|---|---|---|
| Nterm | **Ser Val Leu Cys** Gln Leu Lys Pro Gly Gln Pro Met Tyr | 20 | 0 |
| | **Trp Gln Trp Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | 100 | 10 |
| | **Trp Gln Trp Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | 100 | 10 |
| | **Trp His Trp Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | 100 | 100 |
| | **Trp His Trp Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | 100 | 100 |
| | **Trp His Trp Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | 100 | 70 |
| | **Trp Ser Trp Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | 90 | 0 |
| | **Trp Tyr Trp Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | 90 | 0 |
| Mid5 | Trp His Trp Leu **Arg Leu Gln Pro Gly** Gln Pro Met Tyr | 90 | 70 |
| | Trp His Trp Leu **Arg Leu Gln Ser Gly** Gln Pro Met Tyr | 90 | 70 |
| | Trp His Trp Leu **Arg Leu Gly Ala Gly** Gln Pro Met Tyr | 120 | 80 |
| | Trp His Trp Leu **Arg Leu Ser Ala Gly** Gln Pro Met Tyr | 100 | 70 |
| | Trp His Trp Leu **Cys Leu Asp Ala Gly** Gln Pro Met Tyr | 80 | 0 |
| | Trp His Trp Leu **Gln Leu Ser Ala Gly** Gln Pro Met Tyr | 90 | 50 |
| | Trp His Trp Leu **Leu Leu Arg Pro Gly** Gln Pro Met Tyr | 80 | 50 |
| | Trp His Trp Leu **Ser Leu Asp Ala Gly** Gln Pro Met Tyr | 70 | 20 |
| | Trp His Trp Leu **Ser Leu Arg Ala Gly** Gln Pro Met Tyr | 70 | 50 |
| | Trp His Trp Leu **Ser Leu Gly Ala Gly** Gln Pro Met Tyr | 90 | 50 |
| | Trp His Trp Leu **Ser Leu Ser Pro Gly** Gln Pro Met Tyr | 90 | 80 |
| | Trp His Trp Leu **Ser Leu Tyr Pro Gly** Gln Pro Met Tyr | 80 | 50 |
| | Trp His Trp Leu **Ser Leu Val Pro Gly** Gln Pro Met Tyr | 80 | 50 |
| | Trp His Trp Leu **Thr Leu Met Ala Gly** Gln Pro Met Tyr | 80 | 60 |
| | Trp His Trp Leu **Tyr Leu Arg Pro Gly** Gln Pro Met Tyr | 90 | 60 |
| Cterm | Trp His Trp Leu Gln Leu Thr Pro Gly **Gln Pro Met Tyr** | 100 | 100 |

[a] Peptide sequences are predicted to be encoded by plasmids isolated from the Nterm, Mid5, and Cterm libraries by genetic selection for Ste2p agonists. Amino acids encoded by intentionally variable codons are shown in boldface. Each peptide is encoded by at least one unique nucleotide sequence.
[b] Growth activity in SC-His of CY1455 transformed with peptide-encoding plasmid as a percentage of growth of CY1455 transformed with α-factor-encoding plasmid.
[c] Values shown are from single experiments in which all strains were simultaneously tested and are rounded to the nearest multiple of 10%. Values obtained on retesting of particular strains were within 20 and 25%, respectively, of the values shown for the autocrine growth and paracrine arrest activities.
[d] Width of zone of growth arrest of a lawn of CY1892 around cells of CY1455 transformed with peptide-encoding plasmid as a percentage of that observed with CY1455 transformed with α-factor-encoding plasmid. Values are rounded to the nearest multiple of 10%.

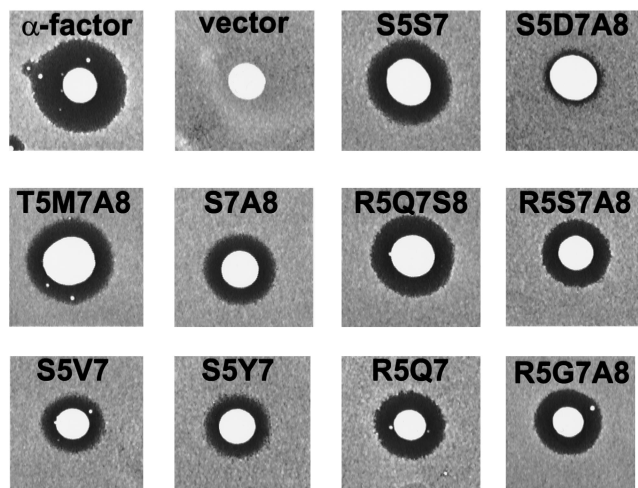


FIG. 5. Paracrine arrest assay. Each panel shows a halo of growth-arrested CY1892 cells surrounding CY1455 cells transformed with a plasmid that expresses genuine α-factor or a Ste2p agonist isolated from the Mid5 library. The peptide predicted to be secreted is indicated on each panel, where S5S7 indicates [Ser5,Ser7]-α-factor, S5D7A8 indicates [Ser5,Asp7,Ala8]-α-factor, etc. Cells transformed with empty vector fail to produce a halo. The widths of the zones of growth arrest due to these and all other Ste2p agonists found in this study are expressed relative to α-factor in Table 1.

evident for any amino acid at position 1, and four of the six possible amino acids are found at this position among the nine antagonists. Similarly, no bias for any particular amino acid at position 3 is evident among the antagonists.

**Ste2p agonists and antagonists isolated from the Mid5 library.** The Mid5 library encodes the α-factor amino acids W, H, W, and L at positions 1, 2, 3, and 4, respectively, and Q, P, M, and Y at residues 10, 11, 12, and 13, respectively. Residues 5 through 9 are unconstrained. Transformation of strain CY1455 with 1 μg of the library yielded $3 \times 10^5$ $Ura^+$ transformants, so that the chance that a particular peptide would have been represented by the CY1455 $Ura^+$ transformants was less than 10%. Genetic selection for Ste2p agonists yielded 16 plasmids that encode tridecapeptides with agonist activity (Table 1); the nucleotide sequences of 2 of these 16 plasmids was identical, suggesting that they were not independent isolates. Statistical analysis indicates the following preferences among the 15 unique predicted peptides. At residue 5 or 7, there is no significant preference for any amino acid. Position 6 shows a strong preference for leucine, with all 15 isolates having a leucine at this position. Position 8 is biased toward proline and alanine. Finally, position 9 shows an apparently absolute preference for glycine. Collectively, the results demonstrate selective pressure for Ste2p agonists to contain a leucine at residue 6, a glycine at position 9, and either a proline or an alanine at position 8. While α-factor conforms to these structural requirements, the apparent ability of alanine to substitute for proline at position 8 is noteworthy in light of the observations of Becker et al. that α-factor forms a type II β turn at Pro8-Gly9 (12–14).

TABLE 2. Genetically selected, pure antagonists of Ste2p

| Library | Predicted amino acid sequence[a] |
|---|---|
| Nterm | **Cys Arg Gly Pro** Gln Leu Lys Pro Gly Gln Pro Met Tyr |
| | **Cys Arg Gly Pro** Gln Leu Lys Pro Gly Gln Pro Met Tyr |
| | **Cys Thr Leu Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr |
| | **Leu His Leu Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr |
| | **Leu His Met Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr |
| | **Phe His Leu Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr |
| | **Trp His Glu Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr |
| | **Trp His Ser Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr |
| | **Trp His Thr Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr |
| | **Trp His Val Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr |
| Mid5 | Trp His Trp Leu **Ala Arg Leu Ala Leu** Gln Pro Met Tyr |
| | Trp His Trp Leu **Arg Leu Trp Gly Leu** Gln Pro Met Tyr |
| | Trp His Trp Leu **Asp Cys Asn Ala Gly** Gln Pro Met Tyr |
| | Trp His Trp Leu **Gln Trp Arg Gln Gly** Gln Pro Met Tyr |
| | Trp His Trp Leu **Gly Leu Gly Pro Ala** Gln Pro Met Tyr |
| | Trp His Trp Leu **Gly Val Ala Val Gly** Gln Pro Met Tyr |
| | Trp His Trp Leu **His Val Gly Ala Gly** Gln Pro Met Tyr |
| | Trp His Trp Leu **Met Leu Arg Pro Gln** Gln Pro Met Tyr |
| | Trp His Trp Leu **Phe Leu Thr Arg Gly** Gln Pro Met Tyr |
| | Trp His Trp Leu **Ser Leu Gly Gly Arg** Gln Pro Met Tyr |
| | Trp His Trp Leu **Ser Leu Gly Arg Leu** Gln Pro Met Tyr |
| | Trp His Trp Leu **Val Leu Gly Asp Leu** Gln Pro Met Tyr |
| Cterm | Trp His Trp Leu Gln Leu Lys Pro Gly **Arg Pro Leu Trp** |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Arg Pro Leu Trp** |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Lys Pro Leu Phe** |

[a] Peptide sequences are predicted to be encoded by plasmids isolated from the Nterm, Mid5, and Cterm libraries by genetic selection for Ste2p antagonists. These plasmids encode no agonist activity. Amino acids encoded by intentionally variable codons are shown in boldface. Each peptide is encoded by at least one unique nucleotide sequence.

The autocrine activities of these agonists are reflected by growth rates of the agonist-expressing strains in liquid media, listed in Table 1. Paracrine measures of agonist activity—induction of growth arrest of CY1892 and stimulation of histidine prototrophy of CY1455—are also listed in Table 1. In the absence of information regarding the production rates of the various agonists, their diffusion rates in solid media, or their permeabilities to the cell wall, these data cannot indicate the relative potencies of the agonists. The data do, however, indicate that paracrine assays of agonist activity may fail to identify agonist peptides identified by the autocrine assay.

Compared with the selection for Ste2p agonists, genetic selection for antagonists from the Mid5 library yielded a greater number of active peptides, as was found with the Nterm library. Transformation of strain CY1892 gave about $2 \times 10^6$ $Ura^+$ transformants, so that the likelihood that a particular peptide would have been among the CY1892 $Ura^+$ transformants was nearly 50%. About 2% of the transformants exhibited α-factor resistance. From 28 transformants chosen for detailed study, 21 exhibited reproducible, plasmid-dependent α-factor resistance. All 21 plasmids isolated from these 21 transformants encode unique tridecapeptides. Twelve of the 21 plasmids failed to confer histidine prototrophy on strain CY1455, indicating that they encode pure antagonists (Table 2). As with the analogs isolated from the Nterm library, there are different profiles of statistically preferred amino acids among the agonists and antagonists isolated from the Mid5 library. Specifically, although there is significant selective pressure for Leu6 among both agonists and antagonists, the preference for Leu6 is greater among the agonists. More dramatically, the conservation of Gly9 that is so pronounced among agonists is not discernible among antagonists. Also, while agonists exhibit a bias for either alanine or proline at position 8, antagonists fail to show any preferences at this position. On the other hand, both agonists and pure antagonists are similar in that they show no preference for any amino acid or class of amino acids at position 5 or 7.

**Ste2p agonists and antagonists isolated from the Cterm library.** The Cterm library is designed to encode the α-factor residues WHWLQLKPG at residues 1 through 9, respectively, of the tridecapeptides, with random amino acids at positions 10 through 13. However, sequence analysis of plasmids isolated from randomly chosen yeast transformants prior to selection for agonist activity indicated greater diversity than intended (see Materials and Methods). Perhaps for this reason, only one peptide with agonist activity was isolated from $8 \times 10^5$ transformants obtained with the Cterm library (Table 1), and this peptide is [Thr7]-α-factor. Genetic selection for antagonists yielded a greater number of active peptides. Of about $2 \times 10^6$ $Ura^+$ transformants of strain CY1892 obtained with the Cterm library, 84 were resistant to α-factor. Detailed study of 21 of these yielded 13 colonies whose plasmids encode 12 different Ste2p antagonists. Strikingly, 10 of the 12 antagonist peptides isolated from the Cterm library exhibit some agonist activity (Table 3), as indicated by the ability of their encoding plasmids to confer autocrine histidine prototrophy when transformed into strain CY1455.

Since only one agonist and two pure antagonists were isolated from the Cterm library, statistical analyses of the C-terminal residues of α-factor are limited. It is clear, however, that each of the three classes of peptides—agonists, antagonists, and partial agonists—exhibits the following numerical preferences: a proline at position 11, a nonaromatic hydrophobic residue at position 12; and an aromatic residue at position 13. Notably, no peptide with agonist activity contains a positively charged residue at position 10, whereas the three pure antagonists contain arginine or lysine at this position.

**Confirmation of agonist and antagonist activities from chemically synthesized peptides.** The amino acid sequences of the bioactive peptides identified in the selections for Ste2p

TABLE 3. Partial Ste2p antagonists

| Library | Predicted amino acid sequence[a] | Autocrine agonist activity[b] |
|---|---|---|
| Nterm | **Phe Asp Tyr Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | + |
| | **Phe His Met Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | + |
| | **Phe His Phe Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | ++ |
| | **Phe His Phe Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | ++ |
| | **Phe His Trp Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | ND |
| | **Ser His Trp Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | +++ |
| | **Trp His Ala Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | + |
| | **Trp His Leu Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | +++ |
| | **Trp His Leu Phe** Gln Leu Lys Pro Gly Gln Pro Met Tyr | + |
| | **Trp His Met Leu** Gln Leu Lys Pro Gly Gln Pro Met Tyr | +++ |
| Mid5 | Trp His Trp Leu **Arg Arg Gly Pro Gly** Gln Pro Met Tyr | ++ |
| | Trp His Trp Leu **Arg Leu Gln Trp Gly** Gln Pro Met Tyr | +++ |
| | Trp His Trp Leu **Arg Leu Gly Arg Met** Gln Pro Met Tyr | ++ |
| | Trp His Trp Leu **Arg Leu Lys Pro Gly** Gln Pro Met Tyr | ++ |
| | Trp His Trp Leu **Gln Val Ser Val Gly** Gln Pro Met Tyr | + |
| | Trp His Trp Leu **Leu Leu Thr Asp Arg** Gln Pro Met Tyr | + |
| | Trp His Trp Leu **Met Leu Leu Ala Gly** Gln Pro Met Tyr | +++ |
| | Trp His Trp Leu **Ser Thr Gly Trp Gly** Gln Pro Met Tyr | ND |
| | Trp His Trp Leu **Thr Leu Val Ser Gly** Gln Pro Met Tyr | +++ |
| Cterm | Trp His Trp Leu Gln Leu Lys Pro Gly **Gln Ala Leu Phe** | ++ |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Gln Gln Gln Trp** | ++ |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Gln Gly Leu Trp** | ++ |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Gln Ser Val Tyr** | ++ |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Glu Ala Met Trp** | ++ |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Glu Ser Leu Tyr** | + |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Met Pro Leu Trp** | ++ |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Ser Gly Leu Tyr** | + |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Ser Pro Leu Trp** | ++ |
| | Trp His Trp Leu Gln Leu Lys Pro Gly **Thr Pro Met Ser** | + |

[a] Peptide sequences are predicted to be encoded by plasmids isolated from the Nterm, Mid5, and Cterm libraries by genetic selection for Ste2p antagonists. These plasmids encode agonist activity, as reflected by their ability to confer autocrine histidine prototrophy upon transformation of CY1455. Amino acids encoded by intentionally variable codons are shown in boldface. Each peptide is encoded by at least one unique nucleotide sequence.

[b] The level of aminotriazole resistance, a measure of autocrine agonist activity, is graded in terms of the maximal aminotriazole concentration on which transformed CY1455 can grow: +, 0.2 mM; ++, 1 mM; +++, 5 mM. For comparison, CY1455 transformed with α-factor-encoding plasmid can grow at aminotriazole concentrations as high as 50 mM. ND, not done.

agonists and antagonists were predicted from the nucleotide sequence of their encoding plasmids. To assess the reliability of these predictions, several peptides were chemically synthesized and tested at approximately 50 to 70% purity, using growth arrest of strain CY1892 as an assay of agonist activity and rescue of CY1892 in the presence of α-factor as an assay of antagonist activity (see Materials and Methods). For instance, [Cys1,Arg2,Gly3,Pro4]-α-factor, which was encoded by two plasmids independently isolated in the genetic screen for Ste2p antagonists, was chemically synthesized and shown to act as a pure antagonist. That is, while the peptide spotted on a lawn of strain CY1892 failed to cause any discernible growth arrest (data not shown), spotting onto an overlay of strain CY1892 on α-factor-containing medium resulted in a zone of rescued growth (Fig. 6). Similarly, chemically synthesized [Ser5,Asp7, Ala8]-α-factor confirmed that this peptide, predicted to have agonist activity, could indeed arrest growth of CY1892 (data not shown). Thus, we conclude that the expression vector used in these studies generally yields the expected peptide sequences and that the autocrine and paracrine effects of the biologically produced peptides reflects their biochemical activities.

Additional peptides were chemically synthesized to test structure-activity inferences made from the collection of agonists and antagonists isolated from each library. For example, Ste2p agonists genetically selected from the Nterm library exhibit significant preferences for Trp1, His2, Trp3, and Leu4, suggesting that α-factor does not readily tolerate mutation of any of its four N-terminal residues. We tested this inference with four chemically synthesized peptides that are point mutations of α-factor (Fig. 7A). Specifically, synthetic [Ser1]-α-factor and [Arg2]-α-factor show reduced though detectable agonist activity, consistent with an important role for Trp1 and His2 for α-factor activity. More dramatically, the virtual ab-

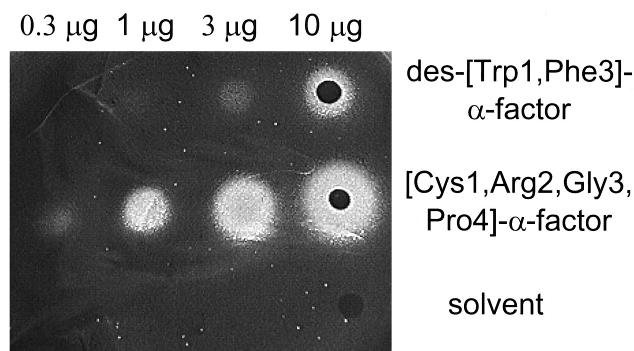


FIG. 6. Antagonist activity of chemically synthesized [Cys1,Arg2,Gly3,Pro4]-α-factor. Peptides were spotted on an overlay of $10^7$ cells CY1892 on solid medium containing 50 nM α-factor. Top row, 1 μl of 10 μg of des-Trp1-[Phe3]-α-factor per μl in dimethyl sulfoxide and serial 1:3 dilutions in water; middle row, 1 μl of 10 μg of [Cys1,Arg2,Gly3,Pro4]-α-factor per μl in dimethyl sulfoxide and serial 1:3 dilutions in water; bottom row, 1 μl of dimethyl sulfoxide and serial 1:3 dilutions in water.

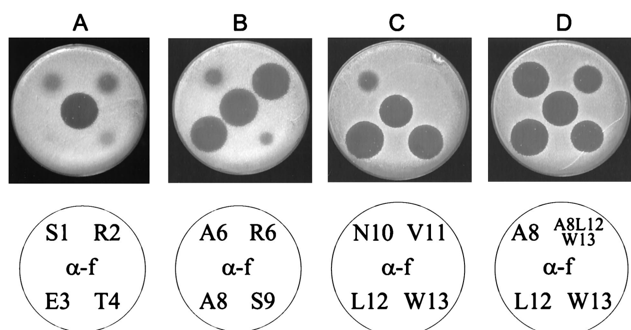


FIG. 7. Agonist activities of chemically synthesized α-factor analogs. Peptides (0.5 μg) were spotted on lawns of $10^7$ cells CY1892. (A) [Ser1]-α-factor (S1), [Arg2]-α-factor (R2), [Glu3]-α-factor (E3), [Thr4]-α-factor (T4), and α-factor (α-f). (B) [Ala6]-α-factor (A6), [Arg6]-α-factor (R6), [Ala8]-α-factor (A8), [Ser9]-α-factor (S9), and α-factor. (C) [Asn10]-α-factor (N10), [Val11]-α-factor (V11), [Leu12]-α-factor (L12), [Trp13]-α-factor (W13), and α-factor. (D) [Ala8]-α-factor (A8), [Ala8,Leu12,Trp13]-α-factor (A8,L12W13), [Leu12]-α-factor (L12), [Trp13]-α-factor (W13), and α-factor.

sence of agonist activity of synthetic [Glu3]-α-factor or [Thr4]-α-factor is consistent with the importance of Trp3 and Leu4 of α-factor for its activity.

Results from the Mid5 library predict that Gln5 and Lys7 are freely mutable, Leu6 and Gly9 are critical for activity, and alanine can substitute for Pro8. All of these predictions except one are confirmed using synthetic peptides (Fig. 7B). As predicted, [Ala8]-α-factor shows agonist activity roughly equivalent to α-factor, as does [Ser5,Asp7,Ala8]-α-factor (data not shown). [Ser9]-α-factor shows virtually no activity, consistent with the importance of Gly9 for α-factor activity. Similarly, the greatly reduced activity of [Ala6]-α-factor is consistent with a critical role for Leu6. Contrary to expectations, however, [Arg6]-α-factor showed growth arrest activity that was equivalent to α-factor. In fact, the equivalence of synthetic [Arg6]-α-factor and α-factor was exceptional, since [Gly6]-α-factor and [Ser6]-α-factor showed, like [Ala6]-α-factor, greatly reduced activity (data not shown). To examine this discrepancy more closely, strain CY1455 was transformed with a plasmid expressing [Arg6]-α-factor or α-factor. Cells transformed with the α-factor-expressing construct showed far greater histidine prototrophy than transformants expressing [Arg6]-α-factor (data not shown). Thus, while chemically synthesized [Arg6]-α-factor is roughly equivalent to genuine α-factor, biologically produced [Arg6]-α-factor has relatively little autocrine activity. This anomaly is discussed below.

The genetic selections of the Cterm library indicate that wild-type residues Gln10 and Pro11 are important for α-factor activity. The reduced activity of synthetic [Asn10]-α-factor and undetectable activity of [Val11]-α-factor (Fig. 7C) are consistent with a requirement for Gln10 and Pro11. Genetic selections also indicate that leucine and tryptophan can substitute for Met12 and Tyr13, respectively. Indeed, synthetic [Leu12]-α-factor and [Trp13]-α-factor are potent Ste2p agonists (Fig. 7C).

Residues other than wild type were more heavily represented at three positions among the genetically selected peptides. That is, alanine, leucine, and tryptophan appeared more frequently than the wild-type proline, methionine, and tyrosine at positions 8, 12, and 13, respectively. Although none of these substitutions, by itself, shows greater agonist activity than genuine α-factor, it is possible that a peptide with all three may show greater activity than α-factor. But chemically synthesized [Ala8,Leu12,Trp13]-α-factor actually shows a smaller zone of growth arrest than either α-factor or any of the single substitutions (Fig. 7D).

## DISCUSSION

We find that Ste2p agonists and antagonists share preferences for histidine, leucine, leucine, proline, a nonaromatic hydrophobic residue, and an aromatic residue at positions 2, 4, 6, 11, 12, and 13, respectively. Agonists show additional preferences for residues tryptophan, tryptophan, proline or alanine, and glycine at positions 1, 3, 8, and 9, respectively (Table 4). We suggest that the preferred residues that are unique to the agonists contribute to receptor activation, while those conserved among both agonists and antagonists participate in Ste2p binding. In the absence of measurements of binding affinity or agonist efficacy, such functional assignments are speculative. Nonetheless, they allow us to hypothesize principal functions for each residue of α-factor that can be tested in future studies (Table 4).

**Limitations of the autocrine selection and data analysis.** We anticipated that not every peptide sequence that could function as an agonist or antagonist would score positive in the auto-

TABLE 4. Interpretations of results and hypotheses of roles of α-factor amino acids

| Position | Functionally important residue[a] | | | Hypothesized role of α-factor residue[b] |
|---|---|---|---|---|
| | Selected agonist | Selected antagonist | α-Factor | |
| 1 | W | | W | Agonist efficacy |
| 2 | H | H | H | Binding affinity |
| 3 | W | | W | Agonist efficacy |
| 4 | L | L | L | Binding affinity |
| 5 | | | Q | ? |
| 6 | L | L | L | Binding affinity |
| 7 | | | K | Cleavage by *BAR1* protease |
| 8 | A or P | | P | Agonist efficacy (structural) |
| 9 | G | | G | Agonist efficacy (structural) |
| 10 | Q | | Q | Agonist efficacy (structural) |
| 11 | P | P | P | Binding affinity |
| 12 | Nonaromatic hydrophobic | Nonaromatic hydrophobic | M | Binding affinity |
| 13 | Aromatic | Aromatic | Y | Binding affinity |

[a] On the basis of results of this study, including genetically selected and chemically synthesized Ste2p agonists and antagonists, we list residues that are preferred among agonists and antagonists. Also listed is the sequence of α-factor, each of whose residues is presumably important for some aspect of α-factor biology.

[b] On the basis of results of this and previous studies (9–14, 18–20, 22, 23, 30), we hypothesize the principal contribution of each amino acid of α-factor to its biology. These hypotheses are speculative and invite experimental testing, bearing in mind that individual residues may serve multiple functions.

crine selection procedure and, reciprocally, that not every tridecapeptide chemically synthesized from the predicted sequence of a positive clone obtained from the selection procedure would have the predicted biological activity. One reason for the lack of complete correspondence could be inappropriate processing of the prepropeptide. The expression vector that we have used is designed to direct the synthesis of a primary translation product containing the MFα prepro sequences appended to the N terminus of the mature peptide. Correct processing of this precursor protein requires endoproteolytic cleavage after the Lys and Arg at residues 84 and 85 of the precursor by Kex2p, followed by removal of two tandem Glu-Ala dipeptides by Ste13p to yield the N terminus of the mature peptide (for a review, see reference 26). At some frequency, Kex2p will cleave within the peptide sequence. Such proteolytic activity may account for the dramatic difference in activities of chemically synthesized and plasmid-encoded [Arg6]-α-factor. That is, substitution of Leu6 with arginine will generate an internal Arg-Lys dipeptide, which may be recognized by Kex2p (7). Also, Ste13p may not always generate the predicted N terminus of the secreted peptide. Although the activity of Ste13p is relatively insensitive to residues following N-terminal Glu-Ala dipeptides (4, 5), some of the peptides encoded by the plasmid libraries may influence Ste13p activity. Other conceivable complications can occur with expression of peptides that fortuitously contain glycosylation signals, stop-transfer signals, etc. In such cases, the active species may not be the peptide that is predicted to be secreted. In general, this can be addressed by testing the effects of exogenously added, chemically synthesized peptides, bearing in mind such confounding issues as permeability of the cell wall to the peptides.

We have interpreted the conservation of particular residues among both agonists and antagonists in terms of direct interaction with the presumed target, Ste2p. However, a conserved residue may be required not for interaction with Ste2p per se but rather for peptide presentation to the receptor. Leu6, for example, may satisfy a constraint in transport or maturation of the peptide (constraints that may not be met by arginine, for example) and may be unimportant for receptor binding. This could explain the substantial bias for Leu6 among agonists and antagonists identified in the genetic screen, despite the nearly wild-type activity of synthetic [Arg6]-α-factor.

We have used peptides chemically synthesized to 50 to 70% purity to test certain results of the autocrine selection procedure. It seems unlikely that similar unintended analogs would be generated during chemical and biological synthesis of a particular peptide. For this reason, results obtained with chemically made peptides are most reliable when they corroborate results obtained with in vivo selection. When the results differ, however, the impurity of the chemically synthesized peptide precludes clear interpretation of the results. For example, the apparently wild-type biological activity of chemically but not biologically synthesized [Arg6]-α-factor may be due to a potent analog contaminating the synthesized [Arg6]-α-factor.

Lastly, our analysis treats each position of the tridecapeptides independently of the 12 others, ignoring possible interactions among positions. In fact, the functions of residues are likely interdependent, admitting the possibility that combinations of mutations may be permitted in agonists while the single mutations are inactive. Our data set is too small to test for such interdependence. With a sufficiently large data set, however, nonrandom coincidence of particular residues at particular positions could be detected, perhaps allowing identification of determinants defined by multiple residues.

**Structure-activity inferences from autocrine selection of agonists and antagonists. (i) The amino-terminal domain.** The results obtained in this study generally confirm a model of α-factor in which the N-terminal region contains major determinants of agonist activity. In particular, the dramatic conservation of Trp1 and Trp3 among agonists but not pure antagonists argues strongly for their importance in activating Ste2p, consistent with the inactivity of chemically synthesized [Ser1]-α-factor and [Glu3]-α-factor. On the other hand, conservation of His2 and Leu4 among both agonists and antagonists indicates that these residues are important for both activities. In the case of Leu4, the findings are consistent with the virtual absence of activity of chemically synthesized [Thr4]-α-factor. In the case of His2, the findings are consistent with the low activity of synthetic [Arg2]-α-factor and with previous studies (3, 19). It is noteworthy that a truncated analog of α-factor that lacks both Trp1 and His2 exhibits about 25% the binding affinity but less than 0.1% the bioactivity of α-factor (10, 22). Our results indicating an important role for His2 for both agonist and antagonist activities may indicate that His2 of α-factor contributes to a spatial positioning of Trp1 that permits interaction with the receptor.

Despite the preference for N-terminal α-factor residues at particular positions among the agonist and antagonist analogs, we recovered both an agonist and an antagonist that differ from α-factor at all of the four N-terminal residues. We chemically synthesized the predicted antagonist, [Cys1,Arg2,Gly3,Pro4]-α-factor, and confirmed it to be an antagonist. Judging from its ability to rescue a larger zone of growth of α-factor-arrested CY1892, [Cys1,Arg2,Gly3,Pro4]-α-factor showed greater antagonist activity than the known inhibitor des-Trp1-[Phe3]-α-factor (22). These peptides indicate that there are multiple solutions to the problems of binding to and, in the case of the agonist, activating Ste2p and that these solutions may be achieved with novel, unanticipated peptides. Notably, both the agonist and the antagonist contain a single cysteine residue, which admits the possibility that the bioactive species is a dimer in both cases.

**(ii) The central domain.** While previous studies have shown that substitution of Leu6 with D-leucine is accompanied by loss of α-factor activity (19), the conservation of Leu6 among all selected peptides indicates its importance for both agonism and antagonism. Pro8 and Gly9, on the other hand, show significant conservation only among agonists, suggesting that they make specific contributions to agonist activity.

Naider and colleagues have obtained substantial evidence that α-factor can form a transient type II β turn centered at Pro8-Gly9 and that this conformation may be important for its activity (11, 12, 30). In view of these findings, the ability of alanine to substitute for proline, suggested by the results of the genetic screen and confirmed with synthetic peptides, implies that alanine may participate in forming a structure that functionally resembles the β turn formed by Pro8-Gly9. Clearly, the apparent functional equivalence of proline and alanine at position 8 is independent of the proposal that α-factor assumes a β turn. Rather, any model that describes the activity of α-factor must account for this equivalence.

Conservative substitutions at positions 5 and 7 have previously been shown to have no effect on the agonist activity of α-factor (22). Our results demonstrate that even coincident, nonconservative substitutions at these positions are tolerated. Thus, [Ser5,Ser7]-α-factor, [Ser5,Val7]-α-factor, [Ser5,Tyr7]-α-factor, [Arg5,Gln7]-α-factor, [Leu5,Gln7]-α-factor, and [Tyr5,Arg7]-α-factor were all identified in the genetic selection as Ste2p agonists. There are, however, limits to the silent substitutions that can occur for Lys7 of α-factor, since its replacement with D-lysine is reported to result in dramatic loss of agonist activity (19). Of course, the relative dispensability of

Gln5 and Lys7 of α-factor is valid only in the narrow context of our selection for Ste2p agonists; these residues may contribute to other biological aspects of α-factor that are not targeted by our selections. For example, α-factor is normally proteolyzed by the *BAR1* protease, which cleaves between Leu6 and Lys7 (9, 23), thereby enabling desensitization of wild-type **a** cells. Thus, one can envision other selection schemes that might demonstrate an important role for Lys7 in the lability of α-factor.

**(iii) The carboxyl domain.** Three aspects of our results are consistent with an important role for C-terminal residues in α-factor activity. First, we find that aromatic residues, especially tryptophan, are heavily overrepresented at position 13 among all selected peptides. The apparent ability of tryptophan to substitute for Tyr13 of α-factor was confirmed by using synthetic [Trp13]-α-factor, which showed agonist activity roughly equivalent to that of α-factor. Second, it has been shown previously that [Nle12]-α-factor, in which Met12 is replaced with the isosteric norleucine, shows nearly wild-type affinity and bioactivity (22). We find that the bioactivity of synthetic [Leu12]-α-factor is roughly equivalent to that of genuine α-factor, suggesting that certain nonsteric substitutions of Met12 are also tolerated. Such nonsteric substitutions must preserve the nonaromatic hydrophobic character of the Met12 side chain. Finally, our results with both genetically selected and chemically synthesized peptides argue that both Gln10 and Pro11 are important for α-factor activity.

**Autocrine genetic selection as a discovery tool.** We have developed an autocrine system that permits genetic selection of Ste2p agonists and antagonists. We have used this system to determine structure-activity relationships for α-factor and, in so doing, have demonstrated its utility for discovering novel activators and inhibitors of Ste2p. Indeed, using a completely random library encoding tridecapeptides, we have recently identified Ste2p ligands that bear no resemblance to α-factor at the primary sequence level. In view of the ability to express mammalian G-protein-coupled receptors in yeast strains (15, 21), this selection system can therefore be used to discover novel peptide agonists and antagonists of known mammalian receptors. The system may also be used to develop surrogate ligands for G-protein-coupled receptors whose physiologic ligands have yet to be identified.

## ACKNOWLEDGMENTS

We thank George Sprague, Ira Herskowitz, and Susan Michaelis for sharing strains; Fred Naider and Y. Larry Zhang for preliminary studies with chemically synthesized agonists; David Fruhling and Michael Spruyt for nucleotide sequencing support; Pedro Sobers for the liquid growth assays; and Lauren Silverman, Jeremy Paul, and Andrew Murphy for helpful discussions.

## REFERENCES

1. **Ammerer, G.** 1983. Expression of genes in yeast using the ADCI promoter. Methods Enzymol. **101:**192–201.
2. **Becker, D. M., and L. Guarente.** 1991. High-efficiency transformation of yeast by electroporation. Methods Enzymol. **194:**182–187.
3. **Blumer, K. J., and J. Thorner.** 1990. β and γ subunits of a yeast guanine nucleotide-binding protein are not essential for membrane association of the α subunit but are required for receptor coupling. Proc. Natl. Acad. Sci. USA **87:**4363–4367.
4. **Boeke, J. D., J. Trueheart, G. Natsoulis, and G. R. Fink.** 1987. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. **154:**164–175.
5. **Bourbonnais, Y., D. Bolin, and D. Shields.** 1988. Secretion of somatostatin by *Saccharomyces cerevisiae*. Correct proteolytic processing of pro-α-factor-somatostatin hybrids requires the products of the *KEX2* and *STE13* genes. J. Biol. Chem. **263:**15342–15347.
6. **Brake, A. J., J. P. Merryweather, D. G. Coit, U. A. Heberlein, F. R. Masiarz, G. T. Mullenbach, M. S. Urdea, P. Valenzuela, and P. J. Barr.** 1984. α-Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **81:**4642–4646.
7. **Brenner, C., and R. S. Fuller.** 1992. Structural and enzymatic characterization of a purified prohormone-processing enzyme: secreted, soluble Kex2 protease. Proc. Natl. Acad. Sci. USA **89:**922–926.
8. **Christianson, T. W., R. S. Sikorski, M. Dante, J. H. Shero, and P. Hieter.** 1992. Multifunctional yeast high-copy-number shuttle vectors. Gene **110:** 119–122.
9. **Ciejek, D., and J. Thorner.** 1979. Recovery of *S. cerevisiae* **a** cells from G1 arrest by α-factor pheromone requires endopeptidase action. Cell **18:**623–635.
10. **Eriotou-Bargiota, E., C.-B. Xue, F. Naider, and J. M. Becker.** 1992. Antagonistic and synergistic peptide analogues of the tridecapeptide mating pheromone of *Saccharomyces cerevisiae*. Biochemistry **31:**551–557.
11. **Garbow, J. R., M. Breslav, O. Antohi, and F. Naider.** 1994. Conformational analysis of the *Saccharomyces cerevisiae* tridecapeptide mating pheromone by $^{13}$C, $^{15}$N rotational-echo double resonance nuclear magnetic resonance spectroscopy. Biochemistry **33:**10094–10099.
12. **Gounarides, J. S., M. S. Broido, J. M. Becker, and F. R. Naider.** 1993. Conformational analysis of [D-Ala$^9$]-α-factor and [L-Ala$^9$]-α-factor in solution in the presence of lipid. Biochemistry **32:**908–917.
13. **Jelicks, L. A., M. S. Broido, J. M. Becker, and F. R. Naider.** 1989. Interaction of the *Saccharomyces cerevisiae* α-factor with phospholipid vesicles as revealed by proton and phosphorus NMR. Biochemistry **28:**4233–4240.
14. **Jelicks, L. A., F. R. Naider, P. Shenbagamurthi, J. M. Becker, and M. S. Broido.** 1988. A type II β-turn in a flexible peptide: proton assignment and conformational analysis of the α-factor from *Saccharomyces cerevisiae* in solution. Biopolymers **27:**431–449.
15. **King, K., H. G. Dohlman, J. Thorner, M. G. Caron, and R. L. Lefkowitz.** 1990. Control of yeast mating signal transduction by a mammalian β2-adrenergic receptor and Gs α subunit. Science **250:**121–123.
16. **Klopotowski, T., and A. Wiater.** 1965. Synergism of aminotriazole and phosphate on the inhibition of yeast imidazoleglycerolphosphate dehydratase. Arch. Biochem. Biophys. **112:**562–566.
17. **Kron, S. J., and N. A. R. Gow.** 1995. Budding yeast morphogenesis: signaling, cytoskeleton and cell cycle. Curr. Opin. Cell Biol. **7:**845–855.
18. **Levin, Y., R. K. Khare, G. Abel, D. Hill, E. Eriotou-Bargiota, J. M. Becker, and F. Naider.** 1993. Histidine of the α-factor of *Saccharomyces cerevisiae* is not essential for binding to its receptor or for biological activity. Biochemistry **32:**8199–8206.
19. **Masui, Y., T. Tanaka, N. Chino, K. Kita, and S. Sakakibara.** 1979. Amino acid substitution of mating factor of *Saccharomyces cerevisiae*. Biochem. Biophys. Res. Commun. **86:**982–987.
20. **Naider, F., J. Gounarides, C.-B. Xue, E. Bargiota, and J. M. Becker.** 1992. Studies on the yeast α-mating factor: a model for mammalian peptide hormones. Biopolymers **32:**335–339.
21. **Price, L. A., E. M. Kajkowski, J. R. Hadcock, B. A. Ozenberger, and M. H. Pausch.** 1995. Functional coupling of a mammalian somatostatin receptor to the yeast pheromone response pathway. Mol. Cell. Biol. **15:**6188–6195.
22. **Raths, S. K., F. Naider, and J. M. Becker.** 1988. Peptide analogues compete with the binding of α-factor to its receptor in *Saccharomyces cerevisiae*. J. Biol. Chem. **263:**17333–17341.
23. **Samokhin, G. P., L. V. Lizlova, J. D. Bespalova, M. I. Titov, and V. N. Smirnov.** 1981. The effect of α-factor on the rate of cell-cycle initiation in Saccharomyces cerevisiae: α-factor modulates transition probability in yeast. Exp. Cell Res. **131:**267–275.
24. **Sapperstein, S., C. Berkower, and S. Michaelis.** 1994. Nucleotide sequence of the yeast *STE14* gene, which encodes farnesylcysteine carboxyl methyltransferase, and demonstration of its essential role in α-factor export. Mol. Cell. Biol. **14:**1438–1449.
25. **Sherman, F.** 1991. Getting started with yeast. Methods Enzymol. **194:**3–21.
26. **Sprague, G. F., Jr., and J. W. Thorner.** 1992. Pheromone response and signal transduction during the mating process of *Saccharomyces cerevisiae*, p. 657–744. *In* E. W. Jones, J. R. Pringle, and J. R. Broach (ed.), The molecular and cellular biology of the yeast *Saccharomyces*: gene expression. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
27. **Stevenson, B. J., N. Rhodes, B. Errede, and G. F. Sprague, Jr.** 1992. Constitutive mutants of the protein kinase STE11 activate the yeast pheromone response pathway in the absence of the G protein. Genes Dev. **6:**1293–1304.
28. **Struhl, K., and R. W. Davis.** 1977. Production of a functional eukaryotic enzyme in *Escherichia coli*: cloning and expression of the yeast structural gene for imidazoleglycerolphosphate dehydratase (his3). Proc. Natl. Acad. Sci. USA **74:**5255–5259.
29. **Taylor, W. R.** 1986. The classification of amino acid conservation. J. Theor. Biol. **119:**205–218.
30. **Xue, C.-B., E. Eriotou-Bargiota, D. Miller, J. M. Becker, and F. Naider.** 1989. A covalently constrained congener of the *Saccharomyces cerevisiae* tridecapeptide mating pheromone is an agonist. J. Biol. Chem. **264:**19161–19168.