MOLECULAR AND CELLULAR BIOLOGY, July 1988, p. 2955–2963
0270-7306/88/072955-09$02.00/0

Vol. 8, No. 7

Genetic Selection for Mutations That Reduce or Abolish Ribosomal Recognition of the *HIS4* Translational Initiator Region

THOMAS F. DONAHUE* AND A. MARK CIGAN

Department of Molecular Biology, Northwestern University Medical School, 303 East Chicago Avenue, Chicago, Illinois 60611

Received 19 February 1988/Accepted 12 April 1988

A unique genetic selection was devised at the *HIS4* locus to address the mechanism of translation initiation in *Saccharomyces cerevisiae* and to probe sequence requirements at the normal translational initiator region that might participate in ribosomal recognition of the AUG start codon. The first AUG codon at the 5′ end of the *HIS4* message serves as the start site for translation, and the −3 and +4 nucleotide positions flanking this AUG (AXXAUGG) correspond to a eucaryotic consensus start region. Despite this similarity, direct selection for mutations that reduce or abolish ribosomal recognition of this region does not provide any insight into the functional nature of flanking nucleotides. The only mutations identified that affected recognition of this region were alterations in the AUG start codon. Among 150 spontaneous isolates, 26 were shown to contain mutations in the AUG start codon, including all +1 changes (CUG, GUG, and UUG), all +3 changes (AUA, AUC, and AUU), and one +2 change (ACG). These seven mutations of the AUG start codon, as well as AAG and AGG constructed in vitro, were assayed for their ability to support *HIS4* expression. No codon other than AUG is physiologically relevant to translation initiation at *HIS4* as determined by growth tests and quantitated in *his4-lacZ* fusion strains. These data and analysis of other *his4* alleles are consistent with a mechanism of initiation at *HIS4* as proposed in the scanning model whereby the first AUG codon nearest the 5′ end of the message serves as the start site for translation and points to the AUG codon in *S. cerevisiae* as an important component for ribosomal recognition of the initiator region.

Ribosomal recognition of the first AUG codon nearest the 5′ end of mRNA is a predominant feature of translation initiation in virtually all eucaryotic organisms (4, 5, 20). In higher eucaryotes, this feature appears to be associated with a ribosomal scanning mechanism whereby ribosomes bind the 5′ end of mRNA and migrate down the message while scanning for a translational start site (17, 18). As part of this model, the first AUG codon in the message has been assumed to act as the signal for ribosomal recognition of the initiator region, since eucaryotic messages appear to lack ribosomal binding sites of the type required for procaryotic translation initiation (17, 36).

Previous genetic and molecular analysis of mutant alleles at the *HIS4* locus in *Saccharomyces cerevisiae* indicated that the first AUG codon nearest the 5′ end of the *HIS4* message serves as the start site for initiation of protein synthesis (9). The *HIS4* initiator region 5′-GAAUAAUG GUU-3′ corresponds in part to both the yeast consensus start region, 5′-A$^{A}_{Y}$A$^{A}_{U}$AAUGUCU-3′ (5), and the vertebrate consensus start region, 5′-CCACCAUGG-3′ (20, 22). Mutational studies of mammalian initiator regions suggest that the sequence context surrounding an AUG start codon is a functional component of the initiator region (16, 21, 22). Simple base changes in nucleotide positions flanking an AUG start site can decrease the efficiency of initiation at an AUG codon. These observations suggest that natural variation of sequence context contributes to establishing alternative rates of initiation at different genes and indicate that the scanning model is more complex than was originally proposed. These mutational studies also show that as a result of decreasing the translational efficiency at an AUG, the ribosome can now bypass this region and initiate at a subsequent downstream AUG. The physiological relevance of this observation would appear to relate to a subset of mammalian messages (<5%) that contain an upstream AUG codon in the 5′ noncoding region that does not serve as the start site for the primary encoded gene product (20). Poor sequence context associated with some of these upstream AUGs (20) might then enable the ribosome to bypass and initiate at a downstream AUG. Therefore, the maintenance of preferred nucleotide context at yeast initiator regions and a similar subset of unusual genes (5%) that have AUGs 5′ to the start of translation (5) suggests that the yeast consensus, although different from the vertebrate consensus, has a similar functional capacity.

As an in vivo attempt to address the mechanism of initiation in *S. cerevisiae* and the functional nature of a translational initiation region, we have used special genetic features, available at the *HIS4* locus, that afford a direct selection for spontaneous mutations with a negative effect on initiation. The unique feature of our selection is that it not only tests the significance of a scanning-type mechanism at *HIS4* but also enables us to directly probe the significance of flanking nucleotide positions that might participate in ribosomal recognition of this region. Our analysis indicates that a scanning-type mechanism is used at *HIS4* for the initiation of translation. However, the only mutations identified that affect recognition of the normal initiator region are alterations of the AUG start codon. This suggests that despite similarities with the yeast and higher eucaryotic consensus, no simple sequence outside the AUG is of great significance to the initiation process at *HIS4*; these results agree with more direct mutational studies in an accompanying report (6).

* Corresponding author.

TABLE 1. *S. cerevisiae* strains used

Strain	Genotype	Source
1-6C	*MATa his4-100 ura3-52 HOL1-1*	T. F. Donahue
E588	*MATα his4-588 ade2-1*	G. R. Fink
4282-1A	*MATα his4-62 ade2*	12
4370-2A	*MATα his4-65 ade2-1*	12
2393A	*MATα his4-280 ade2-1*	G. R. Fink
5470-1C	*MATα his4-712 leu2-3*	7
7082-7D	*MATα his4-864, -1176*	G. R. Fink
A9108B	*MATα his4-24 leu2-3*	12
5194-2B	*MATa his4-24 leu2-3*	12
1050C	*MATα his4-260*	32
17/14	*MATa his1-123*	G. R. Fink
17/17	*MATα his1-123*	G. R. Fink
4-8A	*MATα his4-100 ura3-52 leu2-3, -112 HOL1-1*	T. F. Donahue
45-3B	*MATα his4-401 ura3-52 leu2-3, -112 HOL1-1*	T. F. Donahue
5856-15C	*MATα leu2-3 HOL1-1*	G. R. Fink
TD28	*MATa ura3-52 ino1-13*	8

MATERIALS AND METHODS

Yeast strains, media, and genetic methods. All yeast strains used for this study (Table 1) are ascospore derivatives of *S. cerevisiae* S288C. Strain 1-6C contains a +1 frameshift mutation in the early *HIS4A* coding region between positions +49 and +50 as determined by DNA sequencing (T. F. Donahue, unpublished observation) and was used as the starting strain for the selection of mutations in the initiator codon. Strains were routinely characterized genetically for complementation ability, dominance or recessiveness, and Mendelian segregation. The standard genetic techniques and media used for these studies have been previously described (33). Complementation behavior was tested with the *his4A* mutant strains E588, 4282-1A, and 4370-2A; the *his4C* mutant strains 2393A, 5470-1C, and 7082-7D; the *his4AB* mutant strains A9108B and 5194-2B; the *his4ABC* mutant strain 1050C; and the control strains 17/14 and 17/17. Diploids heterozygous for *HIS4* were selected from crosses with either the His^- strains 4-8A and 45-3B or the His^+ strain 5856-15C to analyze dominance or recessiveness or the segregation of the *HIS4* alleles.

Genetic selection and characterization of revertants. Spontaneous revertants of the polar, *his4A,B,C*, frameshift mutant strain 1-6C were selected by demanding growth on minimal medium plates that contained 3 mM histidinol. Preliminary studies indicated that histidinol-positive (Hol^+) revertants arise at a spontaneous mutation frequency between 5×10^{-8} and 1×10^{-7}. Strain 1-6C was grown in liquid YEPD (yeast extract, peptone, glucose) and plated on YEPD plates which were grown to confluence at 30°C. YEPD plates were then replica plated to synthetic glucose (SD) medium plates containing histidinol and uracil and allowed to incubate at room temperature for approximately 1 week. Within this period, Hol^+ papillae were picked and purified by streaking on YEPD and then retested for *HIS4* expression, as well as for markers indicative of the 1-6C strain. Although not reported, two other strains containing this *his4-100* allele were also studied in reversion tests and gave similar results to those for strain 1-6C; however, we limited our characterization to revertants of 1-6C to maintain the isogenic nature of these spontaneous mutants for comparative growth analysis. All revertants characterized at the molecular level were analyzed genetically in crosses with either $His4^-$ strain 4-8A or 45-3B and the $His4^+$ strain 5856-15C.

Revertant *HIS4* alleles were isolated on a vector plasmid by the integration-excision method (29) as previously described for the isolation of other *his4* alleles (8). The procedures used for preparation of plasmid DNA, restriction mapping, and subcloning, as well as the complete DNA sequence of *HIS4*, have been described previously (9). Mutations were characterized by DNA sequencing. Plasmids were restricted with *Xho*I, and the 5′ end was labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase (24). A second restriction cut with *Pvu*II yields a DNA fragment labeled at only one end. The *Pvu*II-*Xho*I DNA fragment spans positions −588 from the 5′ end of the *HIS4* transcript to +100 in the coding region. This sequencing strategy enables us to confirm the presence of the +1 frameshift mutation between positions +49 and +50 in each revertant and to determine the complete DNA sequence of regions corresponding to the *HIS4* initiation site and 5′ nontranslated portion of the mRNA.

Construction of initiator codon mutant strains. Alleles of *HIS4* were constructed to harbor only the initiator codon mutations in the absence of the +1 frameshift mutation by the transformation/transplacement procedure (14, 31) by using a modification of the method previously described for the construction of other *his4* mutant alleles (8). Unique codon mutations were subcloned as a 767-base-pair (bp) *Sau*3A DNA restriction fragment into the *Bam*HI site of the integrating plasmid YIp5 that contains the selectable yeast marker *URA3* (3). A convenient *Sau*3A site exists at position +34 in the coding region that separates the initiator region from the +1 frameshift. The next *Sau*3A restriction site upstream is at position −670 from the 5′ end of the *HIS4* mRNA. All *HIS4* promoter activities are present in the 5′ region of this *Sau*3A DNA fragment (8, 37). These plasmids were used to transform yeast strain TD28 selecting for $Ura3^+$ function. Transformants were purified by streaking, and enrichment for loss of the vector was performed either by the inositol-less death procedure (13) or the 5′-fluoroorotic acid positive selection method (2) to replace the wild-type allele for the mutated region. His^- Hol^- strains identified by this procedure were purified by streaking and analyzed genetically by the complementation test and in crosses with the $HIS4^+$ strain 5856-15C.

Construction of *HIS4-lacZ* fusion strains. To quantitate the effects of upstream mutations on downstream initiation events, the 767-bp *Sau*3A DNA fragments (as described above) from the wild type and each initiator mutant were subcloned into the unique *Bam*HI restriction site in plasmid p451 (Fig. 1). This plasmid was constructed by ligating the *Hin*dIII-*Sal*I DNA fragment of the *Escherichia coli lac* coding sequences from plasmid pMC1859 (a gift from M. Casadaban) into the *Hin*dIII-*Sal*I fragment of the single-copy *CEN4* vector YCp50 (27). The *Bam*HI site in plasmid p451 is part of a polylinker adaptor at amino acid 8 of the *lacZ* coding region. As a result of inserting the *Sau*3A fragment at this site, a +1 frameshift exists between the *HIS4* and *lacZ* coding regions. However, the AUG codon at position +29 is now in frame with *lacZ* as with our Hol^+ revertant alleles at *HIS4*.

In-frame fusions between the mutant *HIS4* initiator region and the *lacZ* coding region were also constructed to quantitate residual translational expression. The *Sau*3A DNA fragments from the *HIS4* wild type and each initiator mutant were subcloned into the unique *Bam*HI site of plasmid p349 (Fig. 1). This plasmid was constructed to contain the

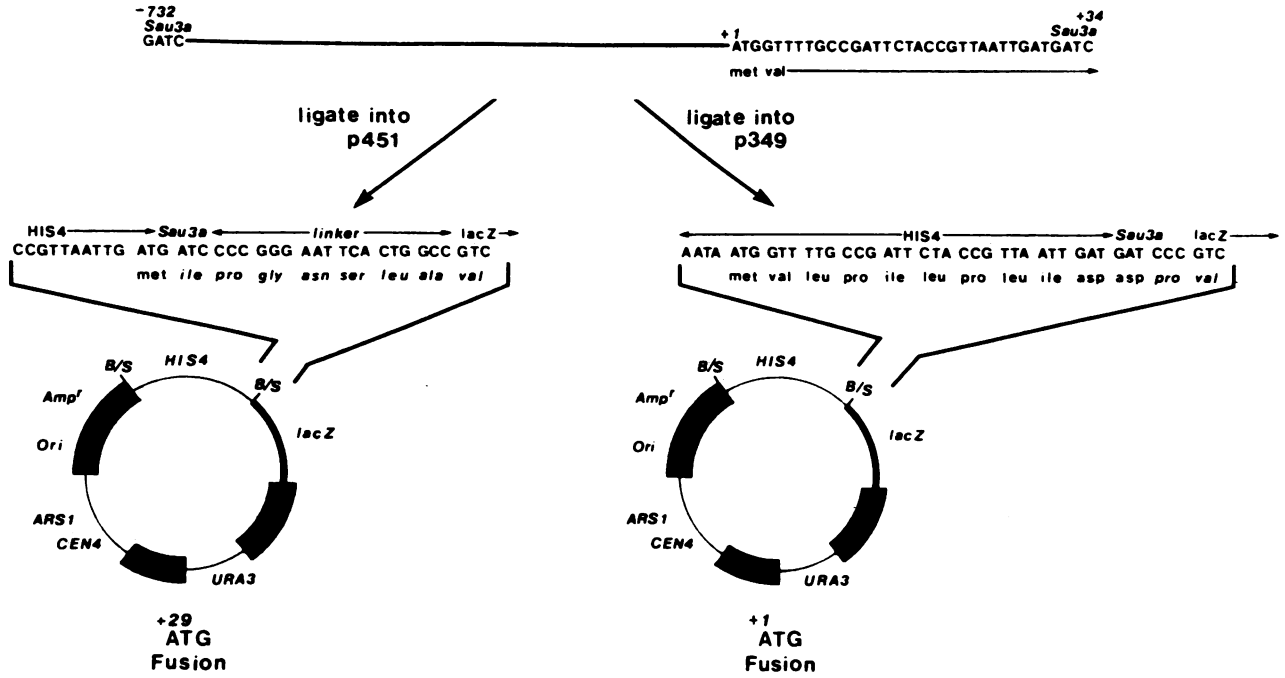

FIG. 1. Construction of *his4-lacZ* fusions. The 767-bp *Sau*3A DNA restriction fragments from wild-type and initiator codon mutant *HIS4* regions were introduced into the *Bam*HI site of the *lacZ*-containing plasmids p451 and p349. The DNA sequence of the *HIS4-lacZ* junction is presented to illustrate the register of both reading frames. Constructions with p451 allow initiation events to be detected from the AUG codon at position +29, while p349 constructions are in frame with the *HIS4* coding region. All plasmids were used to transform *S. cerevisiae* TD28 to quantitate the effects of mutations on downstream initiation (p451 derived) and residual *HIS4* expression (p349 derived). The *Sau*3A DNA fragment from the *HIS4* wild-type region when fused to *lacZ* confers transcriptional and translational properties to *lacZ* identical to those observed at the *HIS4* locus (8). Ori, Origin.

*Bam*HI-*Sal*I DNA fragment of the *lac* coding sequences from plasmid YIp334 (23) ligated into the *Bam*HI-*Sal*I sites of YCp50.

Restriction analysis was performed to confirm the correct orientation of *HIS4-lacZ* constructions. Both p349- and p451-derived constructions were used to transform yeast strain TD28 to Ura3$^+$ by the lithium acetate transformation procedure (15). The wild-type *HIS4 Sau*3A DNA fragment confers expression properties to the β-galactosidase (β-Gal) coding region in *S. cerevisiae* analogous to that of the *HIS4* coding region (37).

β-Gal assays. Yeast extracts were prepared by a modification of the method described by Rose et al. (30). A saturated culture of cells grown at room temperature in supplemented SD medium lacking uracil was used to inoculate 5 ml of the same medium to an optical density at 600 nm of 0.05. Cells were grown overnight to an optical density at 600 nm of 1.2 to 1.3, centrifuged, and suspended in 250 μl of breaking buffer (0.1 M Tris hydrochloride [pH 8.0], 20% [vol/vol] glycerol, 1 mM dithiothreitol) plus 12.5 μl of 40 mM phenylmethylsulfonyl fluoride. All solutions and cells were kept on ice throughout the preparation of the extract. Glass beads (0.45-mm diameter; Sigma Chemical Co.) were added to the meniscus, and the cell suspension was vortexed vigorously for 60 s. Breaking buffer (250 μl) was added to the broken cells, and the solution was centrifuged in a microfuge for 15 min to clarify the extract.

β-Gal assays were performed by the method described by Miller (25). Protein concentrations of yeast extracts were determined by the dye-binding method of Bradford (Bio-Rad Laboratories) with bovine serum albumin as the standard. β-Gal activities were calculated as the number of nanomoles of *o*-nitrophenyl galactoside cleaved per minute per milligram of protein. The specific activity was determined for each extract at three different protein concentrations and was found to be constant. At least two independent yeast strains containing each *HIS4-lacZ* construction were assayed in two independent sets of experiments. The typical specific activity for the in-frame wild-type *HIS4-lacZ* fusion strains was 350 to 450 U. To normalize independent sets of assays, the specific activity of mutant fusion strains is expressed as a percentage of the specific activity obtained for the in-frame wild-type fusion strain included in each set of assays.

In vitro construction of AAG and AGG initiator codon mutations. Oligonucleotide site-directed mutagenesis was used to construct *his4* alleles that contain the mutant initiator codons AAG and AGG. The 17-nucleotide primers 5′-CTGAATAAAGGTTTTGC-3′ and 5′-CTGAATAAGGGTTTTGC-3′ for each construction were synthesized on a Biosearch model 8600 DNA synthesizer with beta-cyanoethyl-diisopropyl phosphoramidite reagents. The template used for mutagenesis was the bacteriophage vector M13mp10 that contained a 1,578-bp *Sal*I DNA restriction fragment from the proximal *HIS4* wild-type region (9). This DNA fragment contains 475 bp of *HIS4* coding sequence, including the translation initiator region, as well as 5′ noncoding sequences from the upstream region of *HIS4*. Single-stranded DNA was isolated and used for mutagenesis (40). The presence of each mutant initiator codon was confirmed by DNA sequencing.

Each respective *Sal*I DNA fragment was ligated into the unique *Sal*I site of the yeast vector YIp5 and used to transform strain TD28 to replace the wild-type *HIS4* region

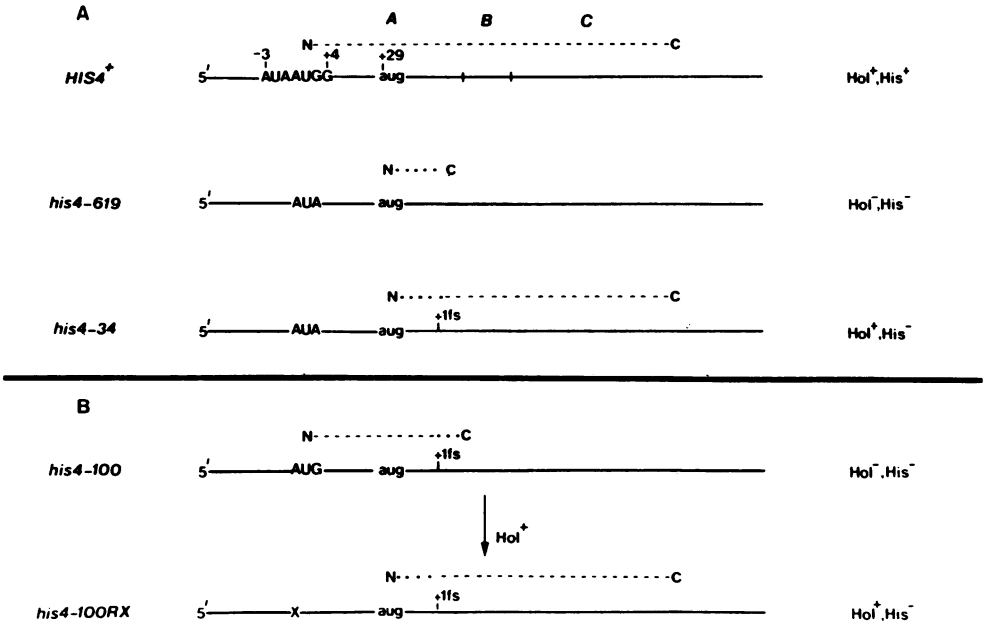

FIG. 2. Strategy of genetic selection scheme. Interpretations of transcription and translation events at wild-type and mutant *HIS4* regions which provide the basis for selection of mutations that reduce or abolish ribosomal recognition of the normal initiator region are schematically presented. (A) The wild-type *HIS4* region is composed of three genetic subregions, A, B, and C, that encode steps 3, 2, and 10 of histidine biosynthesis, respectively. All three enzymatic activities are associated with a single polypeptide (N---C), and initiation begins at the first AUG nearest the 5′ end of the message. The *his4-619* allele contains a point mutation in the first AUG which confers a polar His4$^-$ phenotype. The *his4-34* allele is a revertant of *his4-619* selected by demanding growth on histidinol (Hol$^+$). It contains two mutations: the original G-to-A change at the +3 position and a +1 frameshift between positions +49 and +50. It expresses *HIS4C* (histidinol dehydrogenase) by initiating translation at the first AUG nearest the 5′ end of the message which is now at position +29. This AUG is in the −1 reading frame; however, translation from position +29 (N · · ·) is put back in the correct *HIS4* frame by the downstream +1 frameshift enabling *HIS4C* expression. The DNA sequences of the wild-type, *his4-619*, and *his4-34* alleles have been previously reported (8). (B) The *his4-100* allele was isolated as a polar His$^-$ Hol$^-$ meiotic segregant of a *HIS4*$^+$/*his4-34* diploid and only contains the +1 frameshift mutation of the *his4-34* allele that results in premature termination of translation. Yeast strains containing *his4-100* were reverted to Hol$^+$ to isolate mutations (X) that enable ribosomal bypass of the upstream region as detected by initiation events at the +29 AUG codon that lead to *HIS4C* expression.

with the mutant *his4* region. Mutant strains were genetically analyzed as described above for other initiator codon mutants. The 767-bp *Sau*3A DNA fragment containing either the AAG or AGG initiator codon mutation was ligated into plasmids p349 and p451 and used to construct *HIS4-lacZ* fusion strains as described for the other initiator codon mutations at *HIS4*.

RESULTS

Experimental design. The *HIS4* locus on chromosome III in *S. cerevisiae* encodes a single multifunctional protein which catalyzes three steps in the pathway of histidine biosynthesis (10). The start site for translation is the first AUG codon, 60 nucleotides from the 5′ end of the *HIS4* mRNA (26). The *HIS4* protein is divided into three domains of function, and each of these domains is encoded by a subregion of *HIS4* (11). The *HIS4A*, *B*, and *C* subregions encode, respectively, steps 3 (phosphoribosyl-AMP cyclohydrolase), 2 (phosphoribosyl-ATP pyrophosphohydrolase), and 10 (histidinol dehydrogenase) of histidine biosynthesis. The domains of the protein and the subregions of the gene are colinear with the direction of translation of the polypeptide from A to C (7, 32). Simple missense mutations in one subregion do not affect the expression of others; however, polar mutations in the *HIS4A* region, for example, abolish both *B* and *C* functions. Revertants of polar mutations can be selected either by demanding growth on SD medium lacking histidine (*HIS4ABC*$^+$) or by demanding growth on SD medium supplemented with histidinol (*HIS4ABC*$^+$ or *HIS4C*$^+$) (12). Growth on histidinol (Hol$^+$) requires only expression of the distal *HIS4C* subregion which encodes histidinol dehydrogenase and catalyzes the conversion of histidinol to histidine in the 10th and last step of histidine biosynthesis. The ability to select for distal gene expression independently of total gene expression is an important feature of *HIS4* that makes it particularly suitable for selection of complex mutations.

The genetic and molecular analyses of two mutant alleles of *HIS4* provide the basis for a direct selection of mutations that reduce or abolish ribosomal recognition of the AUG initiator codon (9). The *his4-619* allele is a mutation of the initiator start codon (AUG to AUA) that results in a polar His4A$^-$,B$^-$,C$^-$ phenotype as a result of initiation at the next AUG (position +29) which is out of frame (Fig. 2A). The *his4-34* allele is a revertant of *his4-619* selected for by demanding growth on histidinol (*HIS4C*$^+$). This Hol$^+$ His$^-$ revertant contains two mutations, the original AUG-to-AUA change present in *his4-619* and a second mutation, a +1 frameshift between positions +49 and +50 (Fig. 2A). The explanation for the phenotype of *his4-34* is that the *HIS4C*$^+$ expression is a result of initiation of translation at the AUG at position +29 which is now in frame with the *HIS4*$^+$ reading frame as a result of the distal +1 frameshift. The strain, however, is still His$^-$ as a result of the mutations in the *HIS4A* coding region.

The +1 frameshift mutation from *his4-34* was separated from the AUA mutant codon fortuitously as a meiotic segregant from a *HIS4*$^+$/*his4-34* diploid. This *his4-100* allele

TABLE 2. Compilation of genetic and molecular features of Hol$^+$ His$^-$ revertants

Revertant class	*HIS4C* expression[a]	Total no. of revertants	Nature of mutation	Position of mutation[b]	Strain no.[c]
I	+	2	Deletion	+12 to +27	66
			Deletion	*HIS4A*	86
II	+/−	142	AUG → CUG	+1	34
			AUG → GUG	+1	53, 87
			AUG → UUG	+1	98
			AUG → AUA	+3	24, 36, 38, 46, 47, 49, 65, 111
			AUG → AUC	+3	11, 12, 48, 50, 55, 107, 234, 235
			AUG → AUU	+3	78
			AUG → ACC	+2 and +3	56
			Deletion	+1 to +3	68
III	−/+	6	AUG → ACG	+2	76, 77, 1, 203, 243

[a] As determined by relative growth tests: colony formation on SD plus histidinol plates and ability to genetically complement the *his4AB* mutant yeast strain A9108B on SD-histidine plates.

[b] The A of the AUG start codon is referred to as position +1.

[c] All class I, II, and III strains are spontaneous revertants of the His$^-$ Hol$^-$ parent strain 1-6C.

contained an AUG start codon at the normal initiator site, as confirmed by DNA sequencing. This mutant allele confers a polar Hol$^-$ His$^-$ phenotype, indicating that initiation events at the upstream AUG start codon now preclude initiation events at the downstream AUG codon at position +29 (Fig. 2B). The phenotype of the *his4-100* mutant, compared with that of the *his4-34* mutant, was consistent with predictions of the scanning model whereby initiation events occur at the first AUG codon at the 5′ end of the message. Only in the absence of an upstream initiation site, as in the *his4-34* mutant, could initiation events be detected at a downstream AUG codon. This genetic behavior suggested that the *his4-100* mutant could be used to directly select for mutations that reduce or abolish ribosomal recognition of the first AUG codon at *HIS4* that corresponds to a normal initiator region (Fig. 2B). As with the *his4-34* prototype, the inability to recognize this region should give rise to initiation events at the downstream AUG at position +29. These initiation events result in a Hol$^+$ phenotype as a result of the distal +1 frameshift that restores the correct reading frame and *HIS4C* expression. Revertants of interest should also be His$^-$ because of the defective nature of the *HIS4A* region.

Reversion analysis and characterization of Hol$^+$ revertants. Spontaneous reversion analysis of strain 1-6C was performed in two independent sets of experiments. In the first set, 44 Hol$^+$ His$^-$ revertants were isolated. These revertants were compared in growth tests to assay *HIS4* expression. The basis of this screening was an assumption that some mutations in the upstream region, such as base changes in the AUG, might effectively abolish utilization of this site, thereby maximizing the frequency of initiation events at the next AUG codon, better expression of *HIS4C*, and therefore a stronger Hol$^+$ phenotype. However, other mutations, such as flanking sequences, might only reduce utilization of the upstream AUG start codon, and therefore a weaker Hol$^+$ phenotype may be observed as a result of a lower frequency of initiation events at the downstream AUG. Thus, different Hol$^+$ phenotypes detected among the revertants could directly reflect different types of mutations that might affect ribosomal recognition of this region. One test was simply a comparison of the size of colonies formed by strains streaked out on minimal medium containing histidinol. The second test was genetic complementation to known *his4A*, *B*, and *C* mutants, testing growth ability on minimal medium lacking histidine.

According to these relative growth criteria, Hol$^+$ His$^-$ revertants fell into three distinct classes (Table 2). Class I revertants grew as well on histidinol media as any control *his4A*, *B*, *C*$^+$ mutant did and complemented various *his4A*, *B*, or *C* mutants quite strongly. Two revertants composed this class; one contained a large deletion within the *HIS4A* region as shown by restriction analysis, and the other contained a 16-bp deletion including positions +12 to +27 in the *HIS4* coding region as determined by DNA sequencing. Class II revertants also grew well on histidinol but complemented *his4* mutants less strongly, particularly yeast strain A9108B. Among the 40 revertants in this class, subtle differences were observed in the complementation tests with A9108B. A total of 18 revertants complemented this strain better than the others in this class did. DNA sequencing of seven of these revertant alleles demonstrated simple point mutations in the DNA region corresponding to the normal AUG initiator start codon at *HIS4*, six were of the AUA type, and one encoded GUG at that position. The remaining 22 class II revertants complemented strain A9108B less well than the former 18 did. Of these revertant alleles, 11 were analyzed and shown to encode the following: AUC (5 of 11 revertants); single representatives of CUG, AUU, GUG, and AUA; and two complex mutations, ACC and a precise 3-bp deletion of the AUG. Finally, class III revertants were a minor group which grew poorly on histidinol and barely showed perceivable growth in complementation tests with strain A9108B. Two revertants composed this class, and each allele contained an ACG initiator codon change. Figure 3 illustrates the differences in genetic complementation ability with strain A9108B for representative revertants from classes I, II, and III.

As an attempt to identify other sequences at the *HIS4* initiator region that might participate in ribosomal recognition of this region, we isolated an additional 106 Hol$^+$ His$^-$ revertants and compared them in our growth tests with the first set of revertant strains which had known initiator codon mutations. As in the first series, the majority of revertants (102 of 106) were of the class II type. However, 12 appeared to show subtle growth characteristics unlike the first set of class II revertants. The DNA sequences of 5 of these revertant alleles identified one new initiator codon change, UUG, while the remainder were redundancies of AUC (3 of 5) and AUA (1 of 5). The remaining 4 Hol$^+$ His$^-$ revertants were of the class III type. The DNA sequences of three of these revertant alleles indicated, as before, only the presence of an ACG initiator codon mutation composing class III

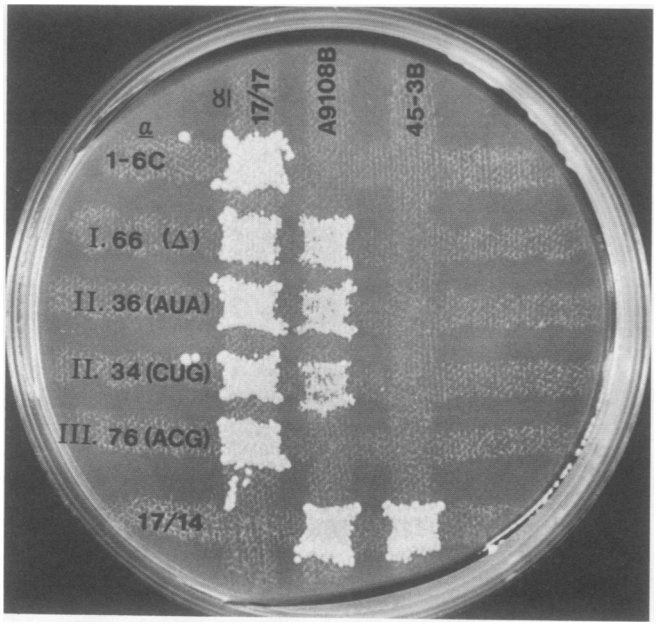

FIG. 3. Differences in genetic complementation ability of representative revertants with strain A9108B. Representative class I, II, and III revertant strains (Table 2) of the **a** mating type are compared for their ability to genetically complement the His4$^-$ *MAT*α yeast strain, A9108B, on SD medium lacking histidine. The *his1* mutant yeast strains 17/14 and 17/17 are included as positive controls for complementation of *his4* mutant yeast strains. The parent strain (1-6C) and *his4* deletion strain (45-3B) are included as negative controls for complementation. *MAT***a** strains are presented horizontally, and *MAT*α strains are presented vertically. Mating that results in diploid strains occurs at the point of intersection of the haploid strains. Growth at the intersection represents a positive complementation response. No growth represents a negative complementation response. The complete genotype of each strain is presented in Table 1, with the exception of revertants which are isogenic derivatives of strain 1-6C.

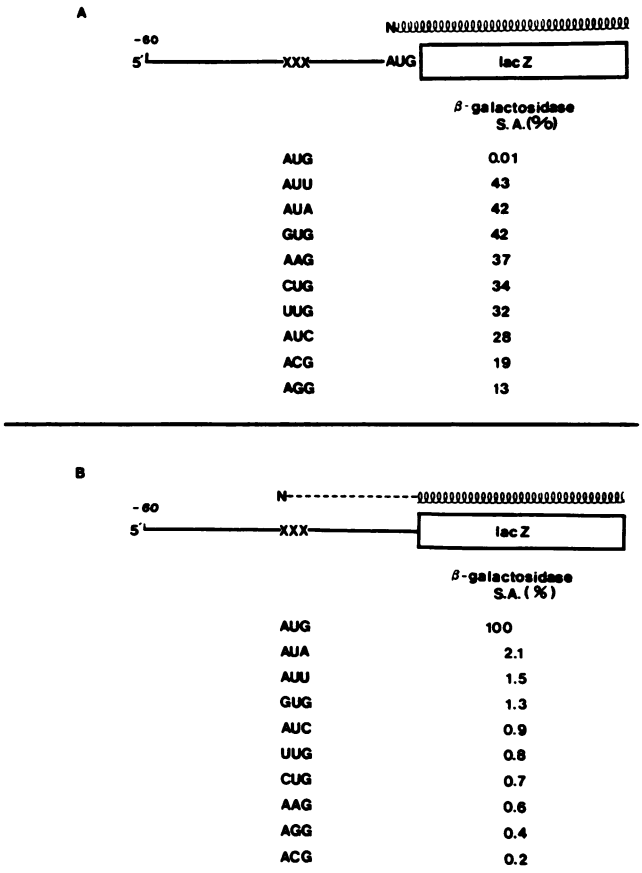

FIG. 4. Effects of initiator codon mutations on *his4-lacZ* expression in *S. cerevisiae*. All nine point mutations of the AUG start codon at *HIS4* were fused as a *Sau*3A DNA fragment to the *lacZ* coding region in either of two reading frames on the YCp50-derived vector p451 and p349 (Fig. 1), and β-Gal activity was assayed for each construction in *S. cerevisiae* TD28 in comparison to a *HIS4*$^+$-*lacZ* fusion control strain. (A) The effects of these mutations on downstream initiation were detected in constructions that enable *lacZ* to be translated (▮▮▮▮▮▮) from the AUG codon at position +29 in the *HIS4* coding region (N). (B) Residual initiation (N---) in the absence of an AUG start codon, as measured by β-Gal activity, was detected in constructions that fuse the *HIS4* coding region in frame with the *lacZ* coding region. XXX, Relative position of the mutant or wild-type AUG triplet in the *HIS4* message (5'——). β-Gal specific activity (S.A.) is expressed as a percentage of the activity obtained for in-frame *HIS4*$^+$-*lacZ* fusion strains to normalize independent sets of assays.

revertants. Therefore, our procedure was capable of distinguishing 7 of the 9 possible point mutations of the AUG start codon (Table 2); however, the procedure was unsuccessful in identifying any functional significance to nucleotide sequences outside the region of the AUG start codon at *HIS4*.

Characterization of initiator codon mutations at *HIS4*. The weak Hol$^+$ phenotype of the class III ACG revertants, together with the inability to identify the AAG and AGG initiator codon mutations, suggested that +2 changes of the AUG start codon, in general, might interfere with the ability to initiate at the +29 AUG that gives rise to *HIS4C* expression. To address this possibility, the AAG and AGG initiator codon mutations were constructed by site-directed mutagenesis and all 9 mutant initiator regions, including the wild-type region, were fused to the *E. coli lacZ* gene such that initiation events at the +29 AUG are in frame with the β-Gal coding region (Fig. 1). These constructions parallel the molecular events at *HIS4* that give rise to *HIS4C* expression in our revertant strains. By measuring β-Gal activities of these two synthesized regions in comparison to the other seven mutant regions, we could determine whether differences in initiation properties at the +29 AUG might be responsible for the absence of AAG and AGG in our selection.

As shown in Fig. 4A, simple base changes at *his4-lacZ* fusions present in these isogenic strains conferred different levels of β-Gal activity. When AUG was present at the normal initiator site, no β-Gal activity was detected, indicating that initiation events at the upstream AUG codon which was out of frame with *lacZ* precluded initiation at the +29 AUG, a genetic behavior identical to that of the *his4-100* allele. When initiator codon mutations were present at the normal initiator site, initiation could be detected at the +29 position. Although the level of β-Gal activities was not comparable to in-frame *HIS4* wild-type fusion controls even when the upstream region was deleted (Donahue, unpublished observation), the relative levels of β-Gal paralleled the relative levels of *HIS4C* expression of class II and class III revertants as indicated by growth criteria (Table 2). In general, mutations at the AUG which conferred higher levels

of *HIS4C* expression to Hol$^+$ class II revertants also conferred higher β-Gal activities, whereas the ACG mutation that conferred lower levels of *HIS4C* expression to Hol$^+$ class III revertants also conferred lower β-Gal activities. These mutational characteristics and the behavior of an AUG codon indicate that effects of simple base changes in the upstream initiator region on β-Gal expression in these *his4-lacZ* fusions mimic the molecular events that give rise to different levels of *HIS4C* expression in our Hol$^+$ revertant strains.

The two initiator codon mutations not identified from our selection and screening procedure had different effects on *lacZ* expression. The AAG +1 frame *his4-lacZ* construction resulted in β-Gal levels similar to those observed for initiator codon mutations derived from class II revertants, and AGG conferred β-Gal activities similar to that of the ACG change. In light of parallels established between *HIS4C* and *lacZ* expression, the higher levels of β-Gal activity associated with the AAG mutation suggest that this mutation should have been identified in reversion studies of the *his4-100* strain. In contrast, the reduced β-Gal activities associated with the +1 frame AGG *his4-lacZ* strain could suggest that this mutation was excluded from our selection because of decreased *HIS4C* expression.

As indicated by the presence of an AUG start codon at the normal initiator site, the inability to initiate at +29 could be a direct consequence of ribosomal recognition of the upstream region that precludes downstream initiation events. The reduced levels of β-Gal and *HIS4C* expression observed when ACG and AGG are present at the normal initiator region could suggest that residual ribosomal recognition of the mutant initiator region precludes initiation at the next available AUG codon. Residual ribosomal recognition of the mutant initiator region may, therefore, be detected as non-AUG initiation events that give rise to reduced levels of gene expression.

We therefore constructed *his4* alleles that only contain each of the nine possible point mutations of the initiator codon and genetically assayed the ability of each codon to support *HIS4* expression. Likewise, each mutant region was fused in frame with *lacZ* to quantitate residual expression in the absence of an AUG start codon (Fig. 1). All initiator codon mutations, when present at *HIS4*, conferred a polar His$^-$ phenotype not complementing either *his4A*, *B*, or *C* mutants. As shown in Fig. 4B, assays of isogenic strains that contain in-frame *his4-lacZ* constructions showed that none of the nine possible point mutations of the normal initiation codon at *HIS4* could support β-Gal activity greater than 2% of control wild-type levels. Specifically, neither the ACG nor AGG initiator codon mutations at *HIS4* gave any indication of perceivable growth on SD medium lacking histidine, even after prolonged periods of incubation (1 week) at either 23, 30, or 37°C, nor did they give any indication of an appreciable level of β-Gal expression.

DISCUSSION

Most of our information about the usage of AUG triplets as initiator codons comes from studies on the *CYC1* gene (the iso-1-cytochrome *c* system) in *S. cerevisiae* (34). A reversion study of mutations in the *CYC1* initiator codon has led to the following empirical conclusions. (i) The AUG at the 5′ end of the message is the site of initiation of translation. (ii) Initiation at this AUG precludes initiation at subsequent AUG codons. (iii) Loss of the first AUG codon by mutation results in initiation at the next AUG codon. Our analysis extends these basic observations at *CYC1* to the *HIS4* locus in *S. cerevisiae*. Two mutant alleles provide the basis for establishing first AUG recognition rules at *HIS4* (9). The *his4-619* allele contains a point mutation in the first AUG codon nearest the 5′ end of the *HIS4* message and confers a polar His4A$^-$, B$^-$, C$^-$ phenotype indicating no residual expression (Fig. 2A). The *his4-34* allele is a revertant of *his4-619* and contains the original mutation at the AUG and a +1 frameshift between positions +49 and +50 (Fig. 2A). This strain is Hol$^+$ but His$^-$ because of mutations in the *HIS4A* region and expresses *HIS4C* activity because of initiation events at the next available AUG nearest the 5′ end of the message at position +29.

That the first AUG codon in the message serves as the start site for translation is further established by our construction of the *his4-100* allele which only contains the +1 frameshift between positions +49 and +50 (Fig. 2B). The presence of an AUG start codon at the normal initiator site in the *his4-100* allele confers a polar His4A$^-$, B$^-$, C$^-$ phenotype, thus precluding initiation events at the +29 AUG that would give rise to *HIS4C* expression. Our inability to detect initiation at the +29 AUG codon in the *his4-100* allele strain indicates that binding of the ribosome at an internal AUG codon does not occur at *HIS4*. Only upon mutating the upstream AUG codon at the normal initiator region, as indicated by our reversion analysis, can initiation at the downstream +29 AUG be detected. This illustrates the 5′-to-3′ progressive nature of translation initiation at AUG codons at *HIS4* and is consistent with predictions of the ribosomal scanning model. These studies, in light of genetic studies of *CYC1* and comparative studies which show that 95% of yeast genes utilize the first AUG codon nearest the 5′ end of the messages as the start site for translation (5), suggest that a scanning-type mechanism exists in *S. cerevisiae* for initiation of protein synthesis.

Our genetic selection scheme at *HIS4* represents the converse genetic approach to reversion studies at *CYC1* (35) and appears to have been an effective method for detecting different and rare mutational events at *HIS4* (Table 2). Classification according to growth defines two different translation initiation events that confer *HIS4C* expression. Class I revertants comprise complex deletions and appear to utilize the normal initiator start codon at the +1 position, as the 16-bp deletion in one revertant allele puts the normal initiator region in frame with *HIS4C*. Class II and III revertants comprise initiator codon mutations and utilize the AUG codon at position +29 as the site of initiation, in agreement with the strategy of the selection scheme. Subtle but discernible growth properties among class II revertants identified eight different alterations of the AUG start codon (CUG, GUG, UUG, AUA, AUC, AUU, ACC, and 3-bp deletion of AUG), whereas class III was unique and distinct from other types of mutations at this position (ACG). The difference between class II and class III revertants is presumably a reflection of the opportunity to utilize the AUG codon at position +29 for initiation.

As a group, base changes at the +2 position of the AUG start codon behaved differently than mutations at either the +1 or +3 position. Although our analysis of +1 frame *his4-lacZ* fusion strains indicates that the AAG mutation should have been isolated as a class II revertant, both ACG and AGG appear to confer reduced levels of translation initiation at the +29 AUG, and for this reason, AGG may have been excluded as a Hol$^+$ revertant. This could reflect residual ribosomal recognition of these upstream mutant initiator regions. It is of interest that an ACG codon has been

implicated to serve as a low-efficiency start codon at a mammalian viral gene (1). However, strains constructed to contain only the ACG codon mutation at *HIS4* do not support any detectable growth on SD medium lacking histidine or any appreciable β-Gal activity in *his4-lacZ* strains. Therefore, if residual recognition of this region does occur, it does not support translation initiation events that are of any significance to *HIS4* expression. In fact, despite the position at the normal initiator region, none of the nine possible point mutations of the AUG start codon at *HIS4* appear to be physiologically relevant as alternative triplets for initiation. Thus, unlike procaryotes which utilize UUG and GUG codons as alternatives to AUG start sites (28), *S. cerevisiae* appears to have a stringent requirement for the AUG triplet as the site of initiation.

In addition to testing the significance of a scanning-type mechanism in *S. cerevisiae*, our selection scheme is unique and enabled us to directly probe sequences other than the AUG codon that might participate in ribosomal recognition of the normal initiator region. A number of genes in *S. cerevisiae*, including *CYC1* and *HIS4*, appear to contain conserved nucleotide positions relative to the AUG start codon, particularly an adenine nucleotide at the −3 position (5); this suggests a potential functional role for these sequences as observed for mammalian messages (19). In addition, recent reports suggest that the −3 position relative to the *CYC1* AUG start codon (39) and an AUG codon within the *CPA1* leader region (38) in *S. cerevisiae* may be functional during the recognition process, although both studies indirectly tested the functional relevance of this position to normal expression patterns at both genes. Our analysis did not identify the functional nature of this position or any other nucleotide position outside the AUG start codon. The possibility exists that mutations at these positions have very subtle effects and that our analysis was insensitive to monitoring these alterations, as our detection procedure was based on initiation events occurring at the AUG codon at position +29, which is not a normal initiator region. The inability to identify simple base changes in conserved nucleotide positions flanking the AUG start codon at *HIS4* could suggest that these sequences may not participate in an equivalent fashion to the AUG codon in recognition of the initiator region. This would agree with earlier reversion studies of initiator codon mutants at *CYC1* (35). The presence of an AUG codon within a 37-nucleotide region was capable of restoring expression of *CYC1*, suggesting that flanking sequences were not essential for initiation.

In a subsequent study, we have directly mutated sequences flanking the AUG start codon at *HIS4* that are shared by the majority of yeast genes (6). These studies support our genetic observation that ribosomal recognition of an initiator region in *S. cerevisiae* may be predominantly controlled by an AUG signal, as opposed to being attenuated by contextual effects as observed with mammalian messages (16, 21, 22).

ACKNOWLEDGMENTS

We thank J. Johansen for excellent technical assistance, R. Gaber and L. Lau for their critical review of this manuscript, Lynne Taylor for its preparation, and G. Fink and C. Styles for numerous yeast strains.

This work was supported by Public Health Service grant GM32263 from the National Institutes of Health to T.F.D.

LITERATURE CITED

1. **Becerra, S. P., J. A. Rose, M. Hardy, B. M. Baroudy, and C. W. Anderson.** 1985. Direct mapping of adeno-associated virus capsid proteins B and C: a possible ACG initiation codon. Proc. Natl. Acad. Sci. USA **82**:7919–7923.
2. **Boeke, J. D., F. LaCroute, and G. R. Fink.** 1984. A positive selection of mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol. Gen. Genet. **197**:345–346.
3. **Botstein, D., S. C. Falco, S. E. Stewart, M. Brennan, S. Scherer, D. T. Stinchcomb, K. Struhl, and R. W. Davis.** 1979. Sterile host yeasts (SHY): a eukaryotic system of biological containment for recombinant DNA experiments. Gene **8**:17–24.
4. **Cavener, D. R.** 1987. Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. Nucleic Acids Res. **15**:1353–1361.
5. **Cigan, A. M., and T. F. Donahue.** 1987. Sequence and structural features associated with translational initiator regions in yeast—a review. Gene **59**:1–18.
6. **Cigan, A. M., E. K. Pabich, and T. F. Donahue.** 1988. Mutational analysis of the *HIS4* translational initiator region in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8**:2964–2975.
7. **Culbertson, M. R., L. Charnas, M. T. Johnson, and G. R. Fink.** 1977. Frameshifts and frameshift suppressors in *Saccharomyces cerevisiae*. Genetics **86**:745–764.
8. **Donahue, T. F., R. S. Daves, G. Lucchini, and G. R. Fink.** 1983. A short nucleotide sequence required for regulation of *HIS4* by the general control system of yeast. Cell **32**:89–98.
9. **Donahue, T. F., P. J. Farabaugh, and G. R. Fink.** 1982. The nucleotide sequence of the *HIS4* region of yeast. Gene **18**:47–59.
10. **Fink, G. R.** 1964. Gene enzyme relations in histidine biosynthesis in yeast. Science **146**:525–527.
11. **Fink, G. R.** 1966. A cluster of genes controlling three enzymes in histidine biosynthesis in *Saccharomyces cerevisiae*. Genetics **53**:445–459.
12. **Fink, G. R., and C. A. Styles.** 1974. Gene conversion of deletions in the *his4* region of yeast. Genetics **77**:231–244.
13. **Henry, S. A., T. F. Donahue, and M. R. Culbertson.** 1975. Selection of spontaneous mutants by inositol starvation in *Saccharomyces cerevisiae*. Mol. Gen. Genet. **143**:5–11.
14. **Hinnen, A., J. A. Hicks, and G. R. Fink.** 1978. Transformation of yeast. Proc. Natl. Acad. Sci. USA **75**:1929–1933.
15. **Ito, H., Y. Fukuda, K. Murata, and A. Kimura.** 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **153**:163–168.
16. **Johansen, H., D. Schumperli, and M. Rosenberg.** 1984. Affecting gene expression by altering the length and sequence of the 5′ leader. Proc. Natl. Acad. Sci. USA **81**:7698–7702.
17. **Kozak, M.** 1978. How do eukaryotic ribosomes select initiation regions in messenger RNA? Cell **15**:1109–1123.
18. **Kozak, M.** 1980. Evaluation of the "scanning model" for initiation of protein synthesis in eucaryotes. Cell **22**:7–8.
19. **Kozak, M.** 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. **9**:5233–5252.
20. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translation start site in eukaryotic mRNAs. Nucleic Acids Res. **12**:857–872.
21. **Kozak, M.** 1984. Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin *in vivo*. Nature (London) **308**:241–246.
22. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44**:283–292.
23. **Lucchini, G., A. G. Hinnebusch, C. Chen, and G. R. Fink.** 1984. Positive regulatory interactions of the *HIS4* gene of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:1326–1333.
24. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
25. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Nagawa, F., and G. R. Fink.** 1985. The relationship between the "TATA" sequence and transcription initiation sites at the *HIS4* gene of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **82**:8557–8561.

27. **Parent, S. A., C. M. Fenimore, and K. A. Bostain.** 1985. Vector systems for the expression, analysis and cloning of DNA sequences in *S. cerevisiae*. Yeast **1**:83–138.
28. **Reddy, P., A. Peterkofsky, and K. McKenney.** 1985. Translational efficiency of the *Escherichia coli* adenylate cyclase gene: mutating the UUG initiation codon to GUG or AUG results in increased gene expression. Proc. Natl. Acad. Sci. USA **82:** 5656–5660.
29. **Roeder, G. S., and G. R. Fink.** 1980. DNA rearrangements associated with a transposable element in yeast. Cell **21**:239–249.
30. **Rose, M., M. J. Casadaban, and D. Botstein.** 1981. Yeast genes fused to β-galactosidase in *Escherichia coli* can be expressed normally in yeast. Proc. Natl. Acad. Sci. USA **78**:2460–2464.
31. **Scherer, S., and R. W. Davis.** 1979. Replacement of chromosome segments with altered DNA sequences constructed *in vitro*. Proc. Natl. Acad. Sci. USA **76**:4951–4955.
32. **Shaffer, B., J. Rytka, and G. R. Fink.** 1969. Nonsense mutations affecting the *his4* enzyme complex of yeast. Proc. Natl. Acad. Sci. USA **63**:1198–1205.
33. **Sherman, F., G. R. Fink, and C. W. Lawrence.** 1972. Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Sherman, F., and J. W. Stewart.** 1982. Mutations altering initiation of translation of yeast iso-1-cytochrome c: contrasts between the eukaryotic and prokaryotic initiation process, p. 301–333. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), The molecular biology of the yeast Saccharomyces. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Sherman, F., J. W. Stewart, and A. M. Schweingruber.** 1980. Mutants of yeast initiating translation of iso-1-cytochrome c within a region spanning 37 nucleotides. Cell **20**:215–222.
36. **Shine, J. and L. Dalgarno.** 1974. The 3′-terminal sequence of *E. coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:** 1342–1346.
37. **Silverman, S. J., M. Rose, D. Botstein, and G. R. Fink.** 1982. Regulation of *HIS4-lacZ* fusions in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **2**:1212–1219.
38. **Werner, M., A. Feller, F. Messenguy, and A. Pierard.** 1987. The leader peptide of yeast gene *CPA1* is essential for the translation repression of its expression. Cell **49**:805–813.
39. **Zitomer, R. S., D. A. Walthall, B. C. Rymond, and C. P. Hollenberg.** 1984. *Saccharomyces cerevisiae* ribosomes recognize non-AUG initiation codons. Mol. Cell. Biol. **4**:1191–1197.
40. **Zoller, M. J., and M. Smith.** 1984. Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. DNA **3**:479–488.