MOLECULAR AND CELLULAR BIOLOGY, Jan. 1994, p. 606–618 Vol. 14, No. 1
0270-7306/94/$04.00+0
Copyright © 1994, American Society for Microbiology

Effect of Sequence Context at Stop Codons on Efficiency of Reinitiation in *GCN4* Translational Control

CHRIS M. GRANT AND ALAN G. HINNEBUSCH*

Section on Molecular Genetics of Lower Eukaryotes, Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, Bethesda, Maryland 20892

Received 28 July 1993/Returned for modification 29 September 1993/Accepted 14 October 1993

Translational control of the *GCN4* gene involves two short open reading frames in the mRNA leader (uORF1 and uORF4) that differ greatly in the ability to allow reinitiation at *GCN4* following their own translation. The low efficiency of reinitiation characteristic of uORF4 can be reconstituted in a hybrid element in which the last codon of uORF1 and 10 nucleotides 3′ to its stop codon (the termination region) are substituted with the corresponding nucleotides from uORF4. To define the features of these 13 nucleotides that determine their effects on reinitiation, we separately randomized the sequence of the third codon and termination region of the uORF1-uORF4 hybrid and selected mutant alleles with the high-level reinitiation that is characteristic of uORF1. The results indicate that many different A+U-rich triplets present at the third codon of uORF1 can overcome the inhibitory effect of the termination region derived from uORF4 on the efficiency of reinitiation at *GCN4*. Efficient reinitiation is not associated with codons specifying a particular amino acid or isoacceptor tRNA. Similarly, we found that a diverse collection of A+U-rich sequences present in the termination region of uORF1 could restore efficient reinitiation at *GCN4* in the presence of the third codon derived from uORF4. To explain these results, we propose that reinitiation can be impaired by stable base pairing between nucleotides flanking the uORF1 stop codon and either the tRNA which pairs with the third codon, the rRNA, or sequences located elsewhere in *GCN4* mRNA. We suggest that these interactions delay the resumption of scanning following peptide chain termination at the uORF and thereby lead to ribosome dissociation from the mRNA.

Regulation of the *GCN4* gene of *Saccharomyces cerevisiae* is one of the best-characterized examples of gene-specific translational control in a eukaryotic organism (reviewed in references 21 and 22). In wild-type yeast cells, *GCN4* mRNA is translated efficiently only when cells are starved for an amino acid. Because GCN4 is a transcriptional activator of amino acid biosynthetic genes, increased translation of *GCN4* mRNA can overcome the amino acid limitation which signalled its derepression. Translational control of *GCN4* depends on four short open reading frames of two or three codons in length present in the *GCN4* mRNA leader, of which the first (uORF1) and fourth (uORF4) are sufficient for nearly wild-type regulation. The presence of uORF4 alone results in constitutively repressed *GCN4* translation. In contrast, uORF1 is a relatively weak translational barrier that must be present upstream of uORF4 for the derepression of *GCN4* translation that occurs in response to amino acid starvation (32).

In addition to the uORFs, derepression of *GCN4* translation in starved cells is dependent on a reduction in the level of active eukaryotic translation initiation factor 2 (eIF-2) (13). The function of eIF-2 in the initiation process is to bind charged initiator $tRNA^{Met}$ in a ternary complex with GTP to the small ribosomal subunit, forming a 43S preinitiation complex (for reviews of translation initiation, see references 35 and 50). The reduction in eIF-2 function that triggers derepression of *GCN4* is mediated by phosphorylation of the α subunit of eIF-2 by the protein kinase GCN2 (13), a mechanism also utilized in mammalian cells to reduce the levels of active eIF-2 under stress conditions (10, 17, 29). According to our recent model (1, 13), ribosomes translate uORF1, remain associated with the mRNA, and resume scanning downstream. Under nonstarvation conditions when the level of active eIF-2 is high, essentially all ribosomes would quickly rebind the eIF-2 · GTP · $Met\text{-}tRNA_i^{Met}$ ternary complex and reinitiate translation at uORF2, -3, or -4, after which they would dissociate from the mRNA. Under starvation conditions when the amount of active eIF-2 is diminished, many ribosomes would fail to rebind the ternary complex until after scanning past uORF2 to -4; those ribosomes which rebind ternary complex while scanning between uORF4 and *GCN4* would be able to reinitiate at *GCN4* instead.

In our model, the positive function of uORF1 resides in its ability to impose a reinitiation mechanism on translation of the downstream ORFs in *GCN4* mRNA (30). It is thought that ribosomes bind the ternary complex before interacting with the 5′ end of mRNA (35). Thus, while reducing the level of eIF-2 activity by phosphorylation of eIF-2 should decrease the rate of ribosome binding at the mRNA cap, it should not cause leaky scanning past the first AUG encountered in an mRNA. This would explain why uORF4 is efficiently recognized under starvation conditions in mutant *GCN4* transcripts when it is the 5′-proximal start site on the mRNA. In contrast to conventional initiation events, ribosomes that translate uORF1 must rebind the ternary complex in the time that it takes to scan from uORF1 to the next AUG codon. As a result of this kinetic constraint, reinitiation at uORF4 can be effectively suppressed by decreasing the concentration of ternary complexes in the cell.

* Corresponding author. Mailing address: NIH, Building 6B, Room 3B-309, Bethesda, MD 20892. Phone: (301) 496-4480. Fax: (301) 496-0243.

The regulatory mechanism just described depends entirely on (i) the tendency of ribosomes to remain associated with the mRNA and resume scanning following translation of uORF1 and (ii) their failure to reinitiate downstream after translating uORF4. We showed previously that moving uORF4 upstream to the position normally occupied by uORF1 decreased the inhibitory effect of uORF4 on *GCN4* expression by two- to threefold (30), in accord with the known dependence of reinitiation on the scanning distance from the uORF to the next start site (28). However, even when uORF4 was present at the position of uORF1 in the *GCN4* leader, uORF4 remained ca. 10-fold more inhibitory than uORF1 to *GCN4* translation. These results suggested that sequence differences at the two uORFs make a large contribution to their very different abilities to allow reinitiation at the *GCN4* start site. This conclusion was strengthened by the fact that replacing the last codon of uORF1 and the 10 nucleotides (nt) immediately following its stop codon with the corresponding nucleotides from uORF4 made uORF1 a strong barrier to *GCN4* translation, equivalent in this respect to uORF4. These same sequence replacements also reduced *GCN4* expression when the mutant copy of uORF1 was situated upstream of uORF4 in the *GCN4* leader. From these results, it was proposed that the wild-type sequences immediately surrounding the termination codon of uORF4 promote ribosome dissociation from the mRNA following translation of the uORF, whereas the wild-type sequences at uORF1 allow ribosomes to resume scanning and reinitiate downstream following translation termination (30).

In the present study, we have undertaken a more extensive analysis of the sequence requirements for efficient reinitiation at *GCN4* following translation of uORF1 in an effort to gain new insight into the process of translation reinitiation. The approach that we took was to randomize sequences of the last triplet of uORF1 and 10 nt following its stop codon and select mutant alleles for which reinitiation occurs at *GCN4* with an efficiency comparable to that given by wild-type uORF1. The results of our analysis indicate that many different sequences rich in A+U bases surrounding the uORF1 termination codon are compatible with efficient reinitiation at *GCN4*. On the basis of these findings, we propose that efficient reinitiation at *GCN4* depends on the ability of ribosomes to exit the uORF1 termination region and resume scanning immediately after peptide chain termination.

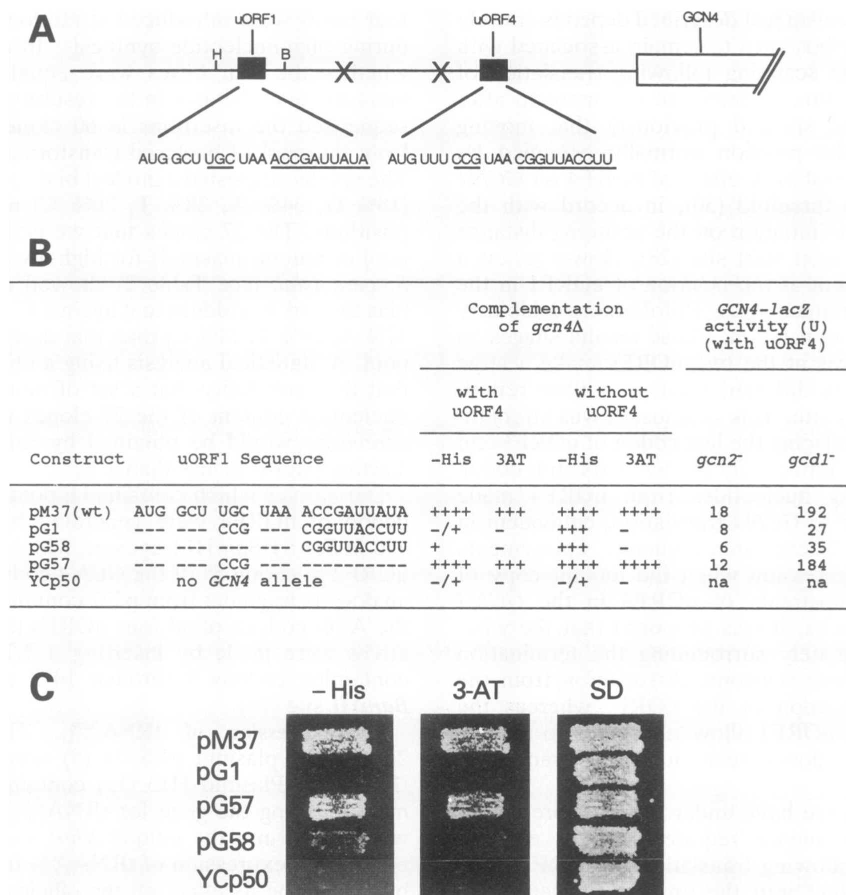
MATERIALS AND METHODS

Construction of mutant *GCN4* alleles. Mutant *GCN4* alleles were generated in plasmid pM37 (30), which is a derivative of the yeast-*Escherichia coli* shuttle vector YCp50 (36), containing the yeast *URA3*, *ARS1*, and *CEN4* sequences for selection and single-copy maintenance in yeast cells. pM37 contains wild-type uORF1 and -4 and point mutations in the ATG codons of uORF2 and -3. In addition, unique *Hin*dIII and *Bgl*II restriction sites flank uORF1 to facilitate cassette mutagenesis of uORF1 and its surrounding sequences (Fig. 1A) (30). Mutagenesis of uORF1 was carried out by insertion of appropriate double-stranded oligonucleotides between the *Hin*dIII and *Bgl*II sites in pM37. The nucleotide sequence of each insertion was verified by dideoxyribonucleotide chain termination sequencing (40). Multiple substitutions were introduced in the 3 nt upstream of the uORF1 stop codon or in the 10 nt immediately downstream of the stop codon by inserting into pM37 mutagenic oligonucleotides in which all four bases were introduced at random in these 13 positions during oligonucleotide synthesis. In an effort to determine whether the four bases were equally represented at the mutagenized positions in the resulting pool of plasmids, we sequenced the insertions in 60 clones selected at random from the pool of bacterial transformants (data not shown). The results suggested a modest bias toward A and T residues (18% G, 34% A, 28% T, 20% C) at the 13 mutagenized positions. The 27 clones that we isolated by screening the pool of mutant plasmids for high-level *GCN4* expression in *S. cerevisiae* (see Table 2) showed a substantially greater bias toward A residues and against C and G residues (9% G, 47% A, 29% T, 14% C) than that determined for the starting pool. A statistical analysis using a chi-square test indicated that the probability that a set of mutant plasmids with the nucleotide content of the 27 clones that we selected in *S. cerevisiae* would be obtained by random sampling of the starting pool was less than 1%.

Derivatives which contained a point mutation in the ATG codon of uORF4 were generated by replacement of the 450-nt *Eco*RI-*Bam*HI fragment, which extends from mutated uORF2 to codon 55 of the *GCN4* coding sequence, with the analogous fragment from p227 containing point mutations in the ATG codons of all four uORFs (32). *GCN4-lacZ* derivatives were made by inserting a 3.2-kb *Bam*HI fragment containing codons 9 through 1023 of *lacZ* at the *GCN4* *Bam*HI site (19).

Overexpression of $tRNA^{Arg}_{CCU}$. The high-copy-number 2μm *TRP1* plasmid pRS424 (8) was used to overexpress $tRNA^{Arg}_{CCU}$. Plasmid H13 (15) contains a 2.5-kb *Xho*I fragment carrying the gene for $tRNA^{Arg}_{CCU}$. This *Xho*I fragment was cloned into the unique *Xho*I site of pRS424, creating pG54. Overexpression of $tRNA^{Arg}_{CCU}$ from pG54 was verified by measuring its effect on the efficiency of frameshifting at the +1 frameshift site of the yeast Ty1 element. Programmed +1 frameshifting in Ty1 requires the sequence CUU-AGG-C and is facilitated by a translational pause at the AGG codon imposed by the low abundance of $tRNA^{Arg}_{CCU}$ (4); accordingly, frameshifting is significantly reduced by overexpression of this tRNA (4, 49). Plasmid pMB38-9merWT contains a +1 frameshift reporter gene consisting of the beginning of *HIS4A* fused to 9 bp containing the Ty1 frameshift site, followed by *lacZ* coding sequences in the −1 frame with respect to *HIS4A*. Thus, expression of β-galactosidase from this construct is a measure of +1 frameshifting in the 9-nt Ty1 sequence. Plasmid pMB38-9merIF is an analogous construct in which *HIS4A* and *lacZ* are fused in frame (4). The frequency of frameshifting in a given strain is calculated as the ratio of β-galactosidase activity produced from construct pMB38-9merWT to that produced from pMB38-9merIF. Overexpression of $tRNA^{Arg}_{CCU}$ from pG54 was verified in our strain by introducing each of the two reporter constructs described above into strains containing pG54 or vector pRS424 alone and measuring β-galactosidase activity in the resulting four strains. The results of this analysis indicated that the presence of pG54 reduced the level of frameshifting from 26% to 5% (data not shown).

Assays of *GCN4* expression. Plasmid-borne *GCN4* alleles were introduced by transformation (25) into strain H1664 (*MAT***a** *ura3-52 leu2-3 leu2-112 trp1-Δ63 his1-29 gcn4*::*LEU2*), and transformants were replica plated to minimal medium (SD) lacking histidine (−His) and SD containing excess (40 mM) leucine and 30 mM 3-aminotriazole (3-AT). *his1-29* is a leaky mutation, and *GCN4*-mediated derepression of *his1-29* is required for growth on −His media. 3-AT is a competitive inhibitor of the *HIS3*-encoded step of the



Construct	uORF1 Sequence	Complementation of *gcn4*Δ, with uORF4: -His	with uORF4: 3AT	without uORF4: -His	without uORF4: 3AT	*GCN4-lacZ* activity (U) (with uORF4): *gcn2*⁻	*gcd1*⁻
pM37(wt)	AUG GCU UGC UAA ACCGAUUAUA	++++	+++	++++	++++	18	192
pG1	--- --- CCG --- CGGUUACCUU	-/+	-	+++	-	8	27
pG58	--- --- --- --- CGGUUACCUU	+	-	+++	-	6	35
pG57	--- --- CCG --- ----------	++++	+++	++++	++++	12	184
YCp50	no *GCN4* allele	-/+	-	-/+	-	na	na

FIG. 1. The sequence of the last codon and termination region of uORF1 affects the efficiency of reinitiation at *GCN4*. (A) Schematic showing the mRNA leader of the *GCN4* allele contained on pM37, with uORF1 and -4 shown as solid boxes and the beginning of the *GCN4* coding region shown as an open rectangle. X, base substitutions in the ATG codons of uORF2 and -3; H and B, *Hin*dIII and *Bgl*II restriction sites used in cassette mutagenesis of uORF1. The last codons and termination regions (10 nt immediately 3′ of the stop codon) of wild-type uORF1 and -4 are underlined. (B) Effects on *GCN4* expression of substituting the last codon or termination region from uORF1 with the corresponding sequences from uORF4. The coding sequence and termination region of uORF1 are given for each construct, with dashes indicating the presence of the same nucleotide present in the parental (wild-type [wt]) construct, pM37. YCp50 is the vector alone, containing no *GCN4* allele. The effects of the mutations were assayed in two different ways. First, *GCN4-lacZ* fusion enzyme activity (expressed as nanomoles of *o*-nitrophenol-β-galactopyranoside hydrolyzed per minute per milligram of protein) was measured in constitutively repressed (*gcn2*⁻) or constitutively derepressed (*gcd1*⁻) mutant strains. Values are averages derived from measurements made on two to three independent transformants from each strain. The individual measurements differed from the mean value by less than 25%. In a second assay, *GCN4* alleles were tested for complementation of the *gcn4*Δ mutation in strain H1664 by measuring the growth of transformants on −His medium with or without the addition of 3-AT. *his1-29* is a leaky mutation, and *GCN4*-mediated derepression of the *his1-29* allele is required for growth on −His medium. 3-AT is a competitive inhibitor of the *HIS3*-encoded step of the histidine biosynthetic pathway and is used to elicit a histidine starvation. The average growth rates of several independently derived transformants examined for each construct are summarized qualitatively by numbers of plus signs. (C) Representative results from the complementation tests summarized in panel B. −His, growth on −His medium after 3 days; 3-AT, growth on 3-AT medium after 6 days; SD, growth on minimal medium supplemented with histidine after 3 days.

histidine biosynthetic pathway that is derepressed by *GCN4* in response to starvation. *GCN4-lacZ* fusion activity was assayed under nonstarvation conditions in strains H15 (*MAT*α *gcn2-1 leu2-3 leu2-112 ura3-52*) and F98 (*MAT*α *gcd1-101 ura3-52*) exactly as described previously (32).

RNA blot hybridization analysis. Total RNA was isolated from cells grown under nonstarvation conditions (SD medium supplemented with histidine, arginine, and tryptophan) and subjected to blot hybridization analysis with radionucleotide probes for *GCN4* and pyruvate kinase (*PYK1*) mRNAs, all as described previously (20) except that a 600-bp *Kpn*I fragment spanning the region from uORF3 to the *Kpn*I site in the *GCN4* coding region was isolated from pG1 and used to prepare the probe for *GCN4* mRNA.

RESULTS

The sequence of the last codon and termination region of uORF1 affects the efficiency of ribosomal reinitiation at *GCN4*. To carry out a mutational analysis of the termination regions of the *GCN4* uORFs, we used the *GCN4* construct in plasmid pM37 (30), containing wild-type uORF1 and -4, point mutations in the ATG codons of uORF2 and -3, and unique restriction sites that allow cassette mutagenesis of

uORF1 and its surrounding sequences (Fig. 1A). uORF2 and -3 are dispensable for regulation of *GCN4* translation, and their absence in pM37 greatly simplifies analysis of the sequence requirements at uORF1 and -4 for translational control. In a previous study (30), the pM37 construct was used to replace sequences surrounding the stop codon at uORF1 with the corresponding nucleotides present at uORF4, and the effects of these mutations on *GCN4* expression were evaluated. By using one assay, the various *GCN4* alleles were tested for the ability to complement a *gcn4Δ* allele in a *gcn4Δ his1-29* strain and restore derepression of genes in the histidine biosynthetic pathway under *GCN4* control. The *his1-29* mutation reduces the activity level of the first enzyme in the histidine biosynthetic pathway, and derepression of *his1-29* transcription by GCN4 is required for growth on −His medium. A more stringent test for the level of *GCN4* expression in *gcn4Δ his1-29* transformants is to demand growth in the presence of 3-AT, an inhibitor of the *HIS3*-encoded step of the histidine pathway. By measuring growth rates of *gcn4Δ his1-29* transformants on −His medium with or without 3-AT, it is possible to measure the expression of different *GCN4* alleles in vivo over a wide range. By using a second assay, *GCN4-lacZ* translational fusions were constructed for each allele, and the amount of β-galactosidase produced by these constructs was measured in *gcn2* and *gcd1* mutant strains that are constitutively repressed or derepressed, respectively, for the general control response. In previous studies, we have observed very good agreement between the results of the complementation tests and the level of *GCN4-lacZ* expression measured in the *gcd1* mutant (1, 30, 32).

The results shown in Fig. 1 summarize and extend the main conclusions of our previous study in showing that replacement of 13 nt surrounding the stop codon of uORF1 with the corresponding sequences from uORF4 severely impairs derepression of *GCN4* expression. As expected, when introduced into the *his1-29 gcn4Δ* strain H1664, the parental construct pM37 allowed growth on both −His and 3-AT media (Fig. 1C and columns labeled "with uORF4" in Fig. 1B), and the corresponding *GCN4-lacZ* fusion showed ca. 11-fold-higher expression in the *gcd1* mutant than in the *gcn2* strain (Fig. 1B). In contrast, the mutant construct in which the third codon and 10 nt from uORF1 were replaced by the corresponding sequences from uORF4 (pG1) conferred poor growth on −His and 3-AT media and greatly reduced *GCN4-lacZ* expression in the *gcd1* strain. The substitutions in the pG1 construct impaired derepression of *GCN4* expression to the same extent as removing the ATG start codon of wild-type uORF1 (32). In constructs pG57 and pG58, the last codon and 10 nt following the uORF1 stop codon, respectively, were replaced with uORF4 sequences to examine the relative contributions of these sequence elements to the inhibition of *GCN4* expression by the uORF1-uORF4 hybrid. Replacing only the 10 bases 3′ to the uORF1 stop codon (henceforth called the termination region) with the corresponding sequences from uORF4 (yielding pG58) diminished derepression of *GCN4* to nearly the same extent as seen for the substitution of all 13 bases in pG1 (Fig. 1). In contrast, substitution of the last codon of uORF1 with the corresponding codon from uORF4 (pG57) had little effect on *GCN4* expression. Introduction of the third codon did, however, appear to exacerbate the inhibitory effect of the uORF4 termination region on *GCN4* expression (Fig. 1; compare pG58 with pG1 for growth on −His medium and *GCN4-lacZ* expression).

A construct that contains wild-type uORF1 alone in the leader exhibits constitutively high levels of *GCN4* expression because uORF4 is required for efficient repression of *GCN4* translation under nonstarvation conditions (31). In accord with our previous results, the substitutions at uORF1 in pG1 and pG58 that reduced the ability of uORF1 to derepress *GCN4* expression in the presence of uORF4 also made uORF1 more inhibitory when it was present alone in the leader (Fig. 1B). Similar findings were interpreted previously to indicate that mutations around the uORF1 stop codon reduce *GCN4* expression because they decrease the frequency of reinitiation following uORF1 translation rather than because they decrease recognition of the uORF1 AUG start codon. If the latter were true, the substitution mutations would be expected to produce higher, not lower, *GCN4* expression from constructs containing uORF1 alone.

A+U-rich codons immediately preceding the stop codon of uORF1 favor reinitiation at *GCN4*. We wished to examine in greater detail the importance of sequences surrounding the stop codons of uORF1 and uORF4 in determining the efficiency of reinitiation at *GCN4*. To do so, we separately randomized the sequences of the third codon or the termination region of the pG1 construct, which contains all 13 bases from uORF4 substituted at uORF1, and isolated *GCN4* alleles that resembled wild-type construct pM37 in complementing the *gcn4Δ* deletion in strain H1664. Plasmids isolated from the His$^+$ 3-AT-resistant (3-AT^r) transformants were shown to confer these phenotypes when reintroduced into strain H1664. Subsequently, DNA sequence analysis was carried out to identify the substitution mutations present in each allele and to confirm the absence of mutations outside of the positions that were randomized in the uORF1 cassette.

Randomization of the triplet preceding the stop codon of the hybrid uORF1-uORF4 sequence in pG1 should generate 64 different alleles, as there are 4^3 different combinations of four bases at three positions. In an effort to cover all 64 possibilities, we screened 242 independent plasmids from the mutagenized pool for complementation of the *gcn4Δ* allele in H1664 and identified 15 different plasmids that conferred a His$^+$ 3-AT^r phenotype comparable to that of pM37 containing wild-type uORF1 sequences. The sequences of the third codons found in the first uORFs of these 15 alleles are shown in Table 1 as triplets without asterisks. At the same time, we characterized plasmids found in three transformants that conferred very poor growth on −His medium and no growth in the presence of 3-AT, the same phenotype seen for the starting pG1 construct containing the third codon from uORF4 and also for pG58 that contains the third codon from uORF1. The *GCN4* alleles in these three transformants contained the triplets UCG, CGG, and CCG (also shown without asterisks in Table 1). Not surprisingly, one of these triplets is the CCG codon found at uORF4. It is noteworthy that despite the minor contribution of the third codon from uORF4 to the inhibitory effect of the uORF1-uORF4 hybrid in pG1 (Fig. 1), we identified 15 codons that could overcome the negative effect of the termination region of uORF4 on ribosomal reinitiation at *GCN4*. The implication of this finding is that the UGC triplet normally present at uORF1 is much less efficient at promoting reinitiation at *GCN4* in the presence of the termination region from uORF4 than are the various triplets that we identified in the 15 highly complementing *GCN4* alleles (Table 1).

We wished to determine the important characteristics of the 15 triplets which allowed them to promote efficient reinitiation at *GCN4* in the presence of the termination region from uORF4. To distinguish between different possi-

TABLE 1. Effects on reinitiation of changing the third codon of uORF1 in the presence of the termination region from uORF4

Construct	3rd Codon of uORF1[a]	Amino acid encoded	Complementation of *gcn4*		Anticodon[b]	RSCU[c]
			−His	3-AT		
pG101	UAA		++++	++	NA	NA
pG3	UGA**		++++	++	NA	NA
pG102	UAU	Tyr	++++	++	GUG	0.132
pG91	AAA	Lys	++++	++	U*UU	0.237
pG103	UCU	Ser	++++	++	IGA	3.359
pG104	UUC	Phe	++++	++	GAA	1.797
pG105	GAA	Glu	++++	++	U*UC	1.968
pG106	UAC	Tyr	++++	++	GUG	1.868
pG107	UUA	Leu	++++	++	U*AA	0.601
pG94	UGU	Cys	++++	++	GCA	1.857
pG109	UUU	Phe	++++	++	GAA	0.203
pG90	AAU	Asn	++++	++	GUU	0.100
pG110	AGA	Arg	++++	++	U*CU	5.241
pG111	CAG	Gln	++++	++	CUG	0.013
pG14	UCA**	Ser	++++	++	U*GA	0.122
pG112	GUA	Val	++++	++	U*AC	0.004
pG95	GUC*	Val	++++	++	IAC	1.796
pG113	AUC	Ile	++++	+	IAU	1.643
pG93	AAG*	Lys	++++	+	CUU	1.763
pG108	CGC*	Arg	++++	+	ICG	0.008
pG99	CUA*	Leu	++++	+	UAG	0.200
pG12	AUA**	Ile	++++	+	U*AU	0.005
pG100	CAA*	Gln	++++	+	U*UG	1.987
pG92	GGU*	Gly	++++	+	ICC	3.898
pG11	UCC**	Ser	+++	−	IGA	2.327
pG96	AAC*	Asn	+++	−	GUU	1.900
pG48	CCU**	Pro	+++	−	IGG	0.179
pG47	CCA**	Pro	++	−	U*GG	3.776
pG41	AGG**	Arg	++	−	CCU	0.017
pG114	UCG	Ser	+	−	CGA	0.170
pG58	UGC** (uORF1)	Cys	+	−	GCA	0.143
pG10	CCC**	Pro	+	−	IGG	0.036
pG84	CGG	Arg	−/+	−	CCG	0.008
pG42	GCG**	Ala	−/+	−	CGC	0.004
pG1	CCG (uORF4)	Pro	−/+	−	CGG	0.009
pG86	GGG**	Gly	−/+	−	CCC	0.017

[a] The sequence of the triplet immediately 5′ of the uORF1 stop codon is shown for each *GCN4* allele tested for its ability to complement the *gcn4Δ* allele in strain H1664 exactly as described for Fig. 1. Triplets without asterisks were identified in *GCN4* alleles selected in *S. cerevisiae* on the basis of conferring wild-type *GCN4* expression. *, triplet obtained by random screening of the pool of mutagenized plasmids contained in bacterial transformants; **, triplet obtained by site-directed mutagenesis.

[b] The predicted anticodons are based on the rules summarized by Guthrie and Abelson (16). NA, not applicable. Asterisks denote base modifications.

[c] RSCU values are from Sharp and Li (42) and reflect the biased usage of synonymous codons.

bilities, it became necessary to characterize triplets in addition to those initially selected from the pool of mutagenized plasmids. Some of these were identified by sequencing 17 plasmids isolated at random from bacterial transformants containing the mutagenized pool (marked with single asterisks in Table 1). Rather than screening a large number of additional plasmids from the mutagenized pool, we decided to introduce the remaining desired triplets by site-directed mutagenesis. The triplets evaluated by this directed approach are indicated in Table 1 by double asterisks.

The first possibility that we considered is that the amino acid encoded by the third triplet may affect the ability of ribosomes to complete elongation, termination, or the resumption of scanning following release of the peptide chain at uORF1. For example, there are several reports indicating that the nascent peptide encoded by an uORF can influence the behavior of the elongating ribosome in a way that affects the flow of scanning ribosomes to the downstream coding sequences (18, 41, 47). The fact that the UAA or UGA nonsense codons present in the third triplet of uORF1 were found to confer high-level reinitiation at *GCN4* in the presence of the uORF4 termination region (Table 1) is consistent with the possibility that reinitiation is impaired by certain amino acids at the C terminus of the uORF1-encoded peptide. However, this hypothesis was made very unlikely by the fact that synonymous codons specifying serine (UCU, UCA, UCC, and UCG), cysteine (UGU and UGC), arginine (AGA, AGG, CGG, and CGC), lysine (AAA and AAG), glutamine (CAA and CAG), and asparagine (AAC and AAU) were found to show markedly different effects on reinitiation (Table 1 and Fig. 2). This is best illustrated by four different arginine codons in Table 1 that each belongs to a different phenotypic group, including the two groups of triplets showing highest or lowest *GCN4* expression.

The second possibility that we considered was that reinitiation would be affected by interactions between the tRNA decoding the third triplet and either the $tRNA^{Ala}_{IGC}$ that decodes the second triplet of uORF1 or with a noncognate tRNA attempting to decode the UAA stop codon of uORF1. Such side-by-side interactions between tRNAs bound to the decoding sites on the ribosome are thought to be important in determining the rate of elongation (14) and the efficiency

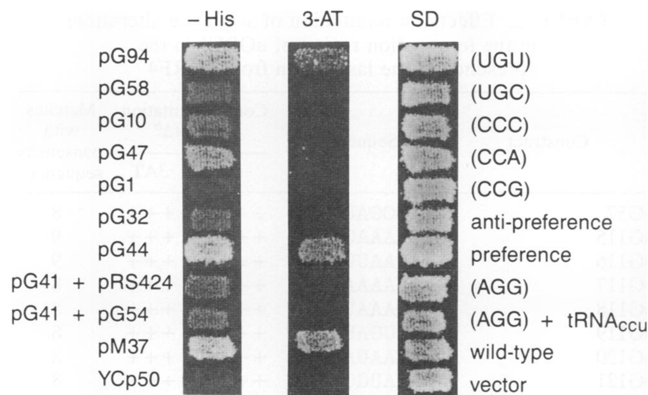


FIG. 2. Expression of selected *GCN4* mutant alleles as measured by complementation of the *gcn4Δ* allele in strain H1664. *GCN4* alleles contained on the indicated plasmids were tested for complementation of the *gcn4Δ* allele in strain H1664 for the inability to grow on –His or 3-AT medium exactly as described for Fig. 1. SD, growth on medium containing histidine. *GCN4* alleles on plasmids pG94, pG58, pG10, pG47, and pG1 contain UGU, UGC, CCC, CCA, and CCG triplets, respectively, at the third codon of uORF1, as shown in Table 1. pG32 and pG44 contain in their termination regions the antipreference and preference sequences, respectively, derived in Table 3. The construct on pG41 has the rare arginine codon AGG at the last codon of uORF1 and is present in strains which were transformed with high-copy-number plasmid pG54 bearing the gene encoding the cognate tRNA$^{Arg}_{CCU}$ or with vector pRS424 alone (see text for details). pM37 contains the wild-type parental *GCN4* allele, and YCp50 is the vector containing no *GCN4* allele.

of nonsense suppression (45). As above, the ability of the two nonsense codons to promote reinitiation at *GCN4* could be explained by this hypothesis, since the presence of these two triplets would eliminate the proposed tRNA-tRNA interactions that hinder reinitiation. This class of hypothesis was made less likely by the fact that the UGU and UGC codons, believed to be decoded by the same cysteine isoacceptor tRNA$^{Cys}_{GCA}$ (16), have very different effects on reinitiation (Table 1 and Fig. 2). Similarly, the serine UCU and UCC codons, the proline CCU and CCC codons, and the asparagine AAC and AAU codons should be decoded by the same isoacceptors (16) but have clearly distinguishable effects on reinitiation at *GCN4* (Table 1).

The third codon in wild-type uORF4 is a rare proline codon (CCG), raising the possibility that codon usage could be correlated with the ability to promote reinitiation. Like most organisms, *S. cerevisiae* shows a biased usage of synonymous codons in highly expressed genes (42, 43), and this bias is correlated with tRNA abundance (24). Rare codons tend to be decoded by the least abundant isoaccepting tRNAs, whereas the most frequently used synonymous codons are recognized by more abundant cognate tRNAs. One way in which the last codon at uORF4 might affect reinitiation would be to cause ribosomes to pause, waiting for a rare tRNA to pair with a nonpreferred codon in the A site of the ribosome. This model could also explain the ability of the stop codons to promote reinitiation as the result of eliminating the elongation pause at the third codon of uORF1. Accordingly, we tabulated the relative synonymous codon usages (RSCU) for each of the codons that were characterized for their effects on reinitiation in Table 1. A high RSCU value indicates that the codon is used much more frequently than the synonymous codons for that amino acid. In general, we found that triplets which disfavor ribosomal reinitiation at *GCN4* had lower RSCU values than the class of triplets which promote reinitiation. There were, however, some notable exceptions to this rule: the GUA valine, AUA isoleucine, and CGC arginine codons have very low RSCU values but were found to promote relatively high levels of reinitiation at *GCN4*, especially for the GUA valine triplet, whereas the CCA proline codon, UCC serine, and AAC asparagine codons have relatively high RSCU values and disfavor continued scanning and reinitiation (Table 1 and Fig. 2).

As an additional test of the idea that triplets disfavor reinitiation at *GCN4* simply because they are decoded by rare tRNAs, we determined the effect on *GCN4* expression of increasing the abundance of the isoacceptor tRNA for the rare arginine codon AGG. This codon was chosen because it has been implicated in the frameshifting mechanism of the yeast Tyl retrotransposons. In these elements, the AGG codon is situated immediately 3′ of the slippery sequence where ribosomes shift reading frame. It is thought that when ribosomes are paused at the AGG codon waiting for the rare tRNA$^{Arg}_{CCU}$ to enter the A site, they have enough time to shift reading frame (4). In accord with this hypothesis, overexpressing tRNA$^{Arg}_{CCU}$ has been shown to reduce the efficiency of frameshifting in Tyl (4, 49). If the AGG codon exerts its negative effect on reinitiation at *GCN4* by making ribosomes pause immediately prior to the uORF1 stop codon, then overexpressing the cognate tRNA$^{Arg}_{CCU}$ should overcome this effect and promote increased expression of *GCN4*.

To test this prediction, we introduced a multicopy plasmid (pG54) encoding tRNA$^{Arg}_{CCU}$ into a transformant of strain H1664 containing the *GCN4* construct with the AGG arginine codon at the third position of the hybrid uORF1-uORF4 element (pG41). This codon is relatively poor in promoting reinitiation at *GCN4* (Table 1). As shown in Fig. 2, the presence of the tRNA$^{Arg}_{CCU}$ gene on the multicopy plasmid had no detectable effect on *GCN4* expression from pG41. We verified that the presence of pG54 in strain H1664 led to a significant reduction in the efficiency of frameshifting in Tyl (from 26% to 5%; see Materials and Methods). This result confirms the idea that being decoded by a low-abundance tRNA is neither necessary nor sufficient to make a particular codon in uORF1 incompatible with reinitiation at *GCN4*.

Finally, we noticed that 6 of the 15 triplets selected by random screening for high levels of reinitiation at *GCN4* contained only A or U bases (AAA, AAU, UUU, UUA, UAU, and UAA). In addition, the remaining codons identified by random and directed mutagenesis which favor reinitiation were predominantly A+U rich, containing only one C or G base. In contrast, the majority of the triplets which conferred low levels of reinitiation at *GCN4* contain C or G bases at either two or all three positions. The correlation between A+U content and efficient reinitiation at *GCN4* for the codons in Table 1 is shown graphically in Fig. 3. On the basis of these findings, we suggest that the defining feature of the many different triplets that confer efficient reinitiation at *GCN4* when present at the third codon of uORF1 is their high A+U content rather than the particular tRNAs with which they pair or the amino acids that they encode.

A+U-rich bases in the termination region of uORF1 favor reinitiation at *GCN4*. To probe the sequence requirements for reinitiation in the uORF1 termination region, we randomized the 10 bases immediately following the stop codon of the uORF1-uORF4 hybrid containing the inhibitory CCG proline codon from uORF4. As above, we selected con-

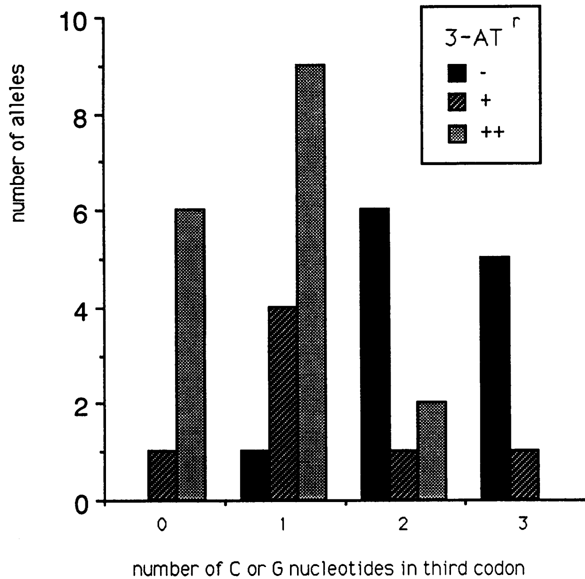


FIG. 3. Efficiency of reinitiation at *GCN4* increases with increasing A+U content in the triplet immediately upstream of the stop codon at uORF1. The *GCN4* alleles shown in Table 1 were analyzed by plotting the qualitative growth rates on 3-AT medium that they conferred on *gcn4Δ* strain H1664, a measure of *GCN4* expression, as a function of the C+G content of the third triplet in uORF1. *GCN4* alleles were assigned to one of three expression classes, conferring either no growth (−), weak growth (+), or strong growth (++) on 3-AT medium.

structs which permit a high level of reinitiation at *GCN4* comparable to that given by the construct containing wild-type uORF1. The alleles that we obtained, derived from a pool of 7,200 different mutant plasmids, represented a small fraction of the 4^{10} possible alleles containing different combinations of four bases at the 10 positions following the uORF1 stop codon. We selected alleles allowing the highest level of *GCN4* expression by plating transformants of strain H1664 (*gcn4Δ his1-29*) bearing mutagenized plasmids directly onto minimal medium supplemented with a low concentration of histidine (4.3 μM). Under these growth conditions, strain H1664 transformed with vector alone or with the hybrid uORF1-uORF4 plasmid pG1 cannot form colonies, whereas introduction of plasmid pG57 containing the uORF1 termination region and the uORF4 CCG proline codon allows efficient colony formation (data not shown). Using this selection technique, we isolated 27 alleles that restored growth on −His and 3-AT media in strain H1664 to roughly the same extent seen for construct pG57. These phenotypes and the sequences of the uORF1 termination regions for the selected alleles are shown in Table 2.

Inspection of the sequences listed in Table 2 revealed a very heterogeneous collection of termination regions that promote efficient reinitiation at *GCN4* (Table 3). Thus, many different sequences can mimic the termination region of uORF1 and promote efficient reinitiation at *GCN4* in the presence of the CCG proline codon from uORF4. Although the termination regions shown in Table 2 are diverse in sequence, significant preferences were observed at most positions. For example, A was found at position 1 (immediately 3′ of the stop codon) in 20 of 27 different alleles, and a G was never found in this position. Certain bases, such as G at positions 3 and 8, were significantly underrepresented at other positions as well. On the basis of the occurrence of bases at each position in the 27 sequences analyzed, we derived a preference sequence containing the most preferred base at each position, which was either A or U. We also determined an antipreference sequence containing the least preferred bases, which were predominantly C or G bases (Table 3). As described in Materials and Methods, the starting pool of mutagenized plasmids appeared to be somewhat biased toward A residues; however, our statistical analysis indicated that the sequence preferences that we detected among the termination sequences that promote efficient reinitiation were very unlikely to have arisen by chance alone (see Materials and Methods).

As a test of the functional significance of the sequence preferences shown in Table 3, we asked whether the derived preference and antipreference sequences would have the expected opposing effects on reinitiation at *GCN4*. To answer this question, we constructed plasmids pG44 and pG32 containing the preference and antipreference sequences,

TABLE 2. Effects on reinitiation of sequence alterations in the termination region of uORF1 in the presence of the last codon from uORF4

Construct	Sequence[a]	Complementation of *gcn4Δ*[b]		Matches with consensus sequence[c]
		−His	3AT	
pG57	ACCGAUUAUA	++++	+++	8
pG115	CUAAAGAAAA	++++	+++	9
pG116	AUAAUUAAAA	++++	+++	9
pG117	UUAAAAUCAU	++++	+++	8
pG118	AGAAAUAAG	++++	+++	8
pG119	ACUGAUUAUA	++++	+++	8
pG120	ACAAUAAACA	++++	+++	8
pG121	AUAUGGUAUA	++++	+++	8
pG122	CUAAUGCAUA	++++	+++	8
pG123	AAAACUAAUA	++++	+++	8
pG124	UGAGAUAAAA	++++	+++	7
pG125	AUAUCGCUUA	++++	+++	7
pG126	AUCUCGAAUU	++++	+++	7
pG127	ACUAGGAUAU	++++	+++	7
pG128	AUAUGAACUU	++++	+++	7
pG129	ACCAAAGAAC	++++	+++	7
pG130	ACUUAUCUUU	++++	+++	7
pG131	ACUAACCCCU	++++	+++	6
pG132	AUAAAUUAUU	++++	++	10
pG133	AUCAAUAAAA	++++	++	9
pG134	ACUAAAGAAG	++++	++	8
pG135	AAAACCAAUA	++++	++	7
pG136	UUACUGUCUU	++++	++	6
pG137	AUCUAAACCA	++++	++	6
pG138	AGUCAAUAUG	++++	++	6
pG139	ACUAAUUUCG	++++	++	6
pG140	UUUGAUCCGA	++++	++	6
pG141	UAAACAAUGU	++++	++	5
pG44 (preference)[d]	AUAAAUAAUA	++++	+++	10
pG32 (antipreference)[d]	GAGCGCGGGC	+	−	0
uORF1[e]	ACCGAUUAUA	++++	+++	8
uORF4[e]	CGGUUACCUU	+	−	4

[a] The 10 nt immediately 3′ to the stop codon of the hybrid uORF1 element containing the CCG triplet from uORF4.

[b] Complementation of the *gcn4Δ* allele in strain H1664 was assessed exactly as described for Fig. 1.

[c] Matches derived in Table 3.

[d] The least and the most preferred nucleotides in the termination region for promoting efficient reinitiation at *GCN4*, derived from Table 3, were used to make plasmids pG32 and pG44, respectively.

[e] The data for uORF1 and uORF4 were obtained from constructs pM37 and pG58, respectively, as shown in Fig. 1.

TABLE 3. Nucleotide frequencies in the termination regions of mutant *GCN4* alleles selected for high-efficiency reinitiation[a]

Determination	Result at indicated position 3′ of stop codon (nt)									
	1	2	3	4	5	6	7	8	9	10
No. of:										
G	0	3	0	3	3	7	2	0	2	4
A	20	3	15	16	15	8	13	16	9	13
U	5	13	8	6	4	10	7	5	12	9
C	2	8	4	2	5	2	5	6	4	1
A+U (%)	93	60	85	81	70	67	75	78	78	81
Consensus	A	U/C	A/U	A	A	A/U/G	A/U/C	A	A/U	A/U
Preference	A	U	A	A	A	U	A	A	U	A
Antipreference	G	A	G	C	G	C	G	G	G	C

[a] The sequences in the termination regions of the mutant *GCN4* alleles listed in Table 2 (constructs pG115 to pG141) were tabulated by determining the number of occurrences of each of the four bases, and the percent A+U, at each nucleotide in the termination region of uORF1. The consensus sequence shows the bases which occur at greater than twice the frequency of the other three bases at that position. When no single base fulfills this criterion, several bases are listed in the consensus, each of which is twice as frequent as the nonconsensus bases at that position. The preference and antipreference sequences contain the most and least frequent bases, respectively, found at each position of the termination region. Note that A and G bases are equally rare at position 2.

respectively, in their termination regions and the third codon from uORF4. As predicted, the preference termination region in pG44 conferred the same high-level *GCN4* expression under starvation conditions seen for the wild-type uORF1 termination region in pG57 and the termination sequences listed in Table 2. Moreover, the antipreference termination region in pG32 behaved much like the uORF4 termination region in pG1 in allowing very little growth under starvation conditions in the *gcn4*Δ mutant (Table 2 and Fig. 2).

On the basis of these findings, we conclude that sequences related to the A+U-rich preference sequence can substitute for the wild-type uORF1 termination region and promote efficient reinitiation at *GCN4* in the presence of the inhibitory CCG proline triplet derived from uORF4. In accord with this conclusion, the termination region of wild-type uORF1 matches the consensus sequence in Table 3 at 8 of 10 positions. This is typical of the sequences listed in Table 2 that confer efficient reinitiation and considerably greater than the 4/10 match shown by the uORF4 termination region (Table 2). It should be noted, however, that the uORF4 termination region is not a good match to the antipreference sequence (Table 2). Thus, similarity to the antipreference sequence in the termination region of uORF1 is not required to impair reinitiation at *GCN4*. It is also noteworthy that the preference termination region is not more efficient at promoting reinitiation at *GCN4* than is the wild-type uORF1 termination sequence or a number of the termination sequences shown in Table 2, despite the presence of several G or C bases in the latter sequences. These results imply that G or C bases can occur at certain combinations of positions in the uORF1 termination region without impairing reinitiation at *GCN4*.

Dissection of the uORF4 termination region. We next sought to determine the relative contributions of bases in the uORF4 termination region to its strong inhibitory effect on reinitiation. Beginning with the pG1 construct (containing the third codon and termination region from uORF4), blocks of nucleotides in the termination region were replaced with the corresponding nucleotides from the preference sequence derived in Table 3. Specifically, the termination region was divided into a block of 3 nt immediately downstream of the stop codon (block 1), a block of 3 nt in the middle of the termination region (block 2), and the final 4 nt at the 3′ end of the termination region (block 3). Replacement of block 1 (CGG) with the preference sequence (AUA) had the strongest effect on reinitiation at GCN4 (Table 4, pG33), restoring good growth on −His medium and weak growth on 3-AT medium. Replacing only block 2 (UUA) with AAU in pG50 had no detectable effect on reinitiation at GCN4, whereas replacing block 3 (CCUU) with AAUA in pG43 restored good growth on −His medium. Thus, replacing blocks 1 and

TABLE 4. Relative contributions of nucleotides from the uORF4 termination region in preventing reinitiation at *GCN4*[a]

Construct	uORF1 sequence	Complementation of *gcn4*Δ	
		−His	3-AT
pG1 (uORF4)	AUG GCU CCG UAA CGGUUACCUU	−/+	−
pG33	—— —— —— —— AUA———	+++	−/+
pG50	—— —— —— —— ——AAU——	−/+	−
pG43	—— —— —— —— ———AAUA	+++	−
pG51	—— —— —— —— AUA——AAUA	++++	++
pG52	—— —— —— —— AUAAAU——	++++	++
pG53	—— —— —— —— ——AAUAAUA	+++	+
pG44	—— —— —— —— AUAAAUAAUA	++++	+++
pG57 (uORF1)	—— —— —— —— ACCGAUUAUA	++++	+++
pG39	—— —— —— —— ——GAUUAUA	++	−

[a] For constructs pG33 to pG44, blocks of nucleotides in the hybrid uORF1-uORF4 element in plasmid pG1 were replaced with the corresponding nucleotides from the preference sequence in Table 3. In constructing pG39, the 3 nt following the stop codon in the termination region of uORF1 (shown in pG57) were replaced with the corresponding nucleotides from the termination region of uORF4 (shown in pG1). Complementation of the *gcn4*Δ allele in strain H1664 was assessed exactly as described for Fig. 1.

TABLE 5. Effects of mutations in the uORF1 termination region on *GCN4* expression in constructs lacking uORF4[a]

Construct	uORF1 sequence	Complementation of *gcn4Δ*			
		With uORF4		Lacking uORF4	
		−His	3-AT	−His	3-AT
pG1	AUG GCU CCG UAA CGGUUACCUU	−/+	−	+++	−
pG57	—— —— —— —— ACCGAUUAUA	++++	+++	++++	++++
pG58	—— —— UGC —— ——————	+	−	+++	−
pG41	—— —— AGG —— ——————	++	−	++++	+
pG48	—— —— CCU —— ——————	++	−	++++	+
pG84	—— —— CGG —— ——————	−/+	−	+++	−
pG42	—— —— GCG —— ——————	−/+	−	+++	−
pG86	—— —— GGG —— ——————	−/+	−	++++	+
pG90	—— —— AAU —— ——————	++++	++	++++	+++
pG91	—— —— AAA —— ——————	++++	++	++++	+++
pG32	—— —— —— —— GAGCGCGGGC	+	−	++++	+
pG38	—— —— UAU —— AUAAAUAAUA	++++	+++	++++	++++
pG33	—— —— —— —— AUA———	+++	−/+	++++	++
pG50	—— —— —— —— ——AAU——	−/+	−	+++	−/+
pG43	—— —— —— —— ————AAUA	+++	−	++++	+
pG51	—— —— —— —— AUA——AAUA	++++	++	++++	+++
pG52	—— —— —— —— AUAAAU——	++++	++	++++	+++
pG53	—— —— —— —— ——AAUAAUA	+++	+	++++	+++
pM37 (wild type)	—— —— UGC —— ACCGAUUAUA	++++	+++	++++	++++

[a] Representative *GCN4* alleles containing the indicated sequences at uORF1 were modified by altering the ATG codon of uORF4 to an ATC codon, thereby eliminating uORF4. The parental *GCN4* alleles containing uORF4 (listed on the left) and the corresponding constructs derived from them which lack uORF4 were analyzed for the ability to complement the *gcn4Δ* allele in strain H1664 exactly as described for Fig. 1.

3 alone led to the greatest increases in *GCN4* expression. Next, pairs of blocks were replaced with the corresponding bases from the preference sequence in plasmids pG51, pG52, and pG53 (Table 4). Substituting the preference bases for blocks 1 and 2, or for blocks 1 and 3, restored high-level reinitiation at *GCN4*, permitting good growth on both −His and 3-AT media. Replacement of blocks 2 and 3 was less effective, restoring good growth on −His plates but only weak growth on 3-AT medium; however, replacing blocks 2 and 3 together (in pG53) did confer higher *GCN4* expression than occurred when these blocks were substituted singly (in pG50 and pG43). In conclusion, blocks 1 and 3 of the uORF4 termination region appeared to make the largest and second-largest contributions, respectively, to the inhibitory effect on reinitiation at *GCN4*. In accord with the idea that high A+U content in the termination region promotes reinitiation, block 1 from uORF4 has the highest G+C content (CGG), followed by block 3 (CCUU), whereas block 2 (UUA) has no C or G bases. On the basis of these results, we predicted that replacing block 1 of the wild-type uORF1 termination region (ACC) with the corresponding bases from uORF4 (CGG) should be sufficient to significantly decrease the efficiency of reinitiation at *GCN4*, but not to the same extent given by introducing the entire uORF4 termination region at uORF1. This prediction was borne out by the results shown in Table 4 for construct pG39, which showed significantly lower *GCN4* expression than the parental construct containing the wild-type uORF1 termination region (pG57) but higher expression than the corresponding construct containing the complete uORF4 termination region (pG1).

Mutations in the uORF1 termination region do not affect initiation at uORF1. To confirm that mutations in the uORF1 termination region affected *GCN4* expression by altering the efficiency of reinitiation rather than by reducing recognition of the uORF1 start codon, a representative set of mutations was examined for their effects on *GCN4* expression in constructs in which uORF1 was present alone in the leader. In accord with previous findings, mutations at uORF1 that impaired derepression of *GCN4* in constructs containing uORF1 and uORF4 also reduced *GCN4* expression in the corresponding constructs containing uORF1 alone (Table 5). If the mutations at uORF1 decreased its ability to overcome the inhibitory effect of uORF4 by causing leaky scanning at the uORF1 AUG codon rather than by reducing reinitiation, such mutations would lead to higher *GCN4* expression in constructs lacking uORF4 because no barrier to ribosomal scanning would remain in the leader. The fact that all of the mutations analyzed lead to lower *GCN4* expression when uORF1 is present alone in the leader indicates that these mutations reduce the ability of ribosomes to reinitiate at *GCN4* following translation of uORF1.

Finally, RNA blot hybridization analysis was conducted to measure the size and steady-state levels of *GCN4* mRNA for a representative set of alleles (Fig. 4). As observed in all

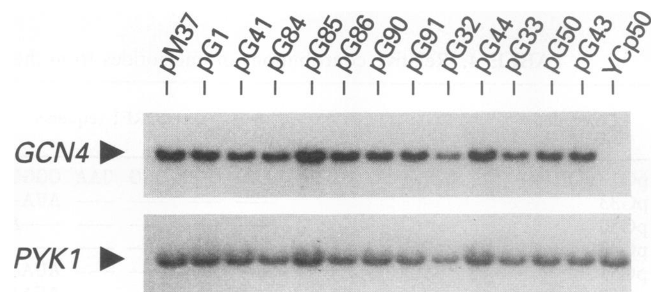


FIG. 4. Steady-state levels of *GCN4* mRNA from selected alleles contained in strain H1664. Transformants of strain H1664 bearing the indicated plasmids were grown to mid-logarithmic phase under nonstarvation conditions. Total RNA was extracted and subjected to RNA blot hybridization analysis using radiolabelled probes for *GCN4* and *PYK1* mRNAs. *PYK1* mRNA was analyzed as a control for the total amount of mRNA loaded in each lane. YCp50 is the vector alone.

previous studies, the mutations at uORF1 were found to have no significant effect on the size or abundance of *GCN4* mRNA (relative to the *PYK1* transcript analyzed in parallel), confirming that they affected *GCN4* expression at the translational level.

DISCUSSION

Previous studies have shown that uORF1 and uORF4 play very different roles in controlling the flow of ribosomes to the *GCN4* start codon. When present as solitary uORFs in their normal positions in the *GCN4* mRNA leader, uORF4 is approximately 30-fold more inhibitory than uORF1 to *GCN4* expression (1, 32). Moving uORF4 upstream into the normal position of uORF1 lessens its inhibitory effect on *GCN4* translation; however, even at the same position in the leader, uORF4 remains about 10-fold more inhibitory than uORF1. Miller and Hinnebusch (30) showed that this difference was conferred primarily by sequences immediately surrounding the stop codons of the two uORFs. For this reason, and because uORF1 and uORF4 were shown to function comparably as translation initiation sites (33), it was suggested that uORF1 is less inhibitory than uORF4 to *GCN4* expression because it allows more ribosomes to resume scanning and reinitiate after its own translation has terminated. The present study was aimed at defining more precisely the sequence requirements at uORF1 for reinitiation at *GCN4* as a first step toward understanding the role of these sequences in translation termination and subsequent reinitiation events.

The last codon of uORF1 could affect the flow of scanning ribosomes to *GCN4* in any of several distinct ways. The ability to promote a high efficiency of reinitiation was conferred by many different triplets specifying numerous amino acids. The triplets that conferred low-level reinitiation also encoded a variety of amino acids. Clearly, the distinct reinitiation properties of uORF1 and uORF4 are not uniquely dependent on the UGC cysteine and CCG proline triplets found in the wild-type uORFs. In fact, we discovered that the UGC triplet in wild-type uORF1 is much less efficient than numerous A+U-rich codons in promoting reinitiation. These observations, combined with the fact that synonymous codons for the same amino acid were found to have markedly different effects on reinitiation, led us to conclude that the particular amino acid present at the C terminus of the nascent uORF1 peptide has no role in determining the reinitiation efficiency. The fact that most triplets that we identified as incompatible with reinitiation are underutilized in yeast genes and probably decoded by low-abundance tRNAs suggested the possibility that the rate of decoding the triplet during the elongation phase of translation will affect the probability of reinitiation. We do not favor this hypothesis, however, because we found that increasing the decoding rate for the AGG arginine triplet by increasing the gene dosage of its cognate tRNA did not have a perceptible effect on reinitiation when the AGG triplet was present in uORF1. In addition, we found that certain poorly utilized codons were compatible with very efficient reinitiation, whereas some preferred codons had a very low reinitiation potential.

Having rejected the idea that the third triplet exerts its effect on *GCN4* expression at the elongation phase of translation, we next considered the possibility that it influences reinitiation by affecting the rate of translation termination and the resumption of scanning following release of the peptide chain. When a ribosome reaches a termination codon (UAA, UAG, or UGA), a peptidyl-tRNA occupies the ribosomal P site and the A site is empty. In eukaryotes, a single release factor (RF) recognizes any of the three termination codons and binds together with GTP to the ribosome. Hydrolysis of the peptidyl-tRNA is catalyzed by the peptidyltransferase, and the release of the peptide chain from the ribosome is followed by GTP hydrolysis. This promotes liberation of the RF from the ribosome and release of the deacylated tRNA (reviewed in reference 11). In *E. coli*, an additional protein known as ribosome-releasing factor (RRF), acting in conjunction with elongation factor G, stimulates dissociation of the ribosome from the mRNA following release of the peptide chain (23, 39). It is not known whether an RRF-like activity is required in eukaryotes. There is evidence, however, that some aspect of the termination stage of protein synthesis in eukaryotes is slow compared with the decoding of sense codons (48). In the model presented below, the third triplet and termination region of uORF1 influence the probability of reinitiation by affecting the final steps of termination and the resumption of scanning following release of the peptide chain.

Our hypothesis is based in part on the observation that *E. coli* ribosomes can reinitiate translation more efficiently downstream of a termination codon in cell-free translation systems depleted of RRF (39). We assume that an RRF-like factor is present in eukaryotes and that it cannot stimulate dissociation of ribosomes from the mRNA once the ribosome (or 40S subunit) returns to the scanning mode following peptide chain release. Thus, the probability of RRF-catalyzed dissociation of the ribosome following release of the peptide chain would be directly proportional to the time that it takes to resume scanning and exit the termination region. Accordingly, a rapid resumption of scanning would increase the probability of reinitiation downstream. We propose that the presence of G+C-rich sequences surrounding the uORF1 stop codon prevents a rapid resumption of scanning following peptide chain release, leading to ribosome dissociation from the mRNA and a low frequency of reinitiation.

The major interactions stabilizing RNA secondary structure are base pairing and stacking. A consideration of thermodynamic parameters for base pairs in RNA secondary structures suggests that interactions involving a CG base pair are more stable than those involving AU base pairs and that the most stable interactions involve adjacent CG base pairs as a result of high stacking energies (44). The marked preference for AU-rich sequences in both the preceding triplet and the 10 nt following the stop codon at uORF1 may indicate that stable base-pairing interactions involving these sequences disfavor completion of the termination process and resumption of scanning. Consideration of the codon substitutions analyzed in Table 1 shows that the codons most restrictive for reinitiation are predicted to form adjacent CG base pairs when pairing with the anticodons of their cognate tRNAs. Thus, one possibility is that strong codon-anticodon interactions at the third codon of uORF1 will lead to slow dissociation of the deacylated tRNA left in the P site of the ribosome following peptide chain release. If dissociation of deacylated tRNA from the P site is required for the resumption of scanning, as seems likely, this would increase the time spent by the ribosome in the termination region and the probability of RRF-catalyzed ribosome dissociation from the mRNA. This notion is consistent with previous experiments showing that purified tRNAs bind with different stabilities to synonymous codons that differ at the wobble position (46). In addition, the stability of the tRNA-mRNA interaction has been demonstrated to affect the frequency of frameshifting

on the RF2 programmed frameshift site in *E. coli* (12). However, given the involvement of inosine and other unusual bases in the anticodons of different tRNAs (reviewed in reference 38) and uncertainty about the contributions made by these nucleotides to base-stacking interactions, the relative stabilities of different codon-anticodon interactions cannot be assigned accurately with the thermodynamic parameters available for predicting RNA secondary structure.

An alternative explanation for the inhibitory effect on reinitiation of G+C-rich triplets at the third codon of uORF1 is that such sequences form stable base pairs with one or more sites on the rRNA. This model could also explain the preference for A+U-rich sequences downstream of the uORF1 stop codon, as the ribosome is thought to cover 30 to 35 nt of mRNA (26, 48). A related explanation for the effects of the sequences 3′ to the stop codon is base pairing with one or more sites on *GCN4* mRNA. Scanning ribosomes are sensitive to secondary structure (27, 37), and it has been shown that introduction of stem-loop structures or G+C-rich sequences into yeast mRNA leaders can dramatically reduce translation of the downstream coding sequences (3, 9). Thus, stable interactions between the termination region of uORF1 and sites either on the rRNA or elsewhere in *GCN4* mRNA could prevent rapid exit of the ribosome from the termination region, increasing the probability of ribosome dissociation from the mRNA.

It could be argued that the sequences surrounding the uORF4 stop codon form a secondary structure that simply blocks the movement of scanning ribosomes to *GCN4* without promoting ribosome dissociation from the mRNA. This model seems unlikely in view of the small size of the uORF4 termination region and the independent inhibitory effects on reinitiation contributed by the last codon and two separate blocks of nucleotides in the termination region. In addition, these sequences do not inhibit *GCN4* translation when they are present at uORF4 in the absence of the uORF4 start codon (20, 32). The fact that the uORF4 sequences reduce *GCN4* expression only when they surround the stop codon of a functional uORF argues that their inhibitory effects on reinitiation are coupled to the termination process. In contrast, insertions of sequences with the potential to form stable stem-loop structures have been shown to inhibit *GCN4* translation independent of their position relative to uORF4 (1).

A final possibility is that the sequences surrounding the stop codon are recognized by the putative RRF or by a sequence in the rRNA that contributes to the termination process and that G+C-rich sequences in the termination region interfere with the function of these termination factors. For example, in *E. coli*, the 3′ end of the 16S rRNA has been implicated in termination and may interact directly with the stop codon (34). However, this region is not conserved in yeast rRNA. The 10 nt downstream of the uORF4 stop codon are complementary in 8 of 10 positions to a sequence present at the 3′ end of 18S rRNA (30), and it has been proposed that this region of the 18S rRNA interacts with the 5S rRNA located in the large subunit (2). This 18S-5S RNA interaction may be involved in the reversible association of the subunits during protein synthesis. However, point mutations which would disrupt crucial base pair interactions in this proposed uORF4 termination region-18S rRNA interaction were found to have no discernible effect on reinitiation (data not shown).

A recent compilation and statistical analysis of sequences flanking translation termination sites revealed no strong consensus sequence for termination in any organism examined; however, certain bases are preferred at the nucleotide immediately 3′ to the stop codon (5–7). In *S. cerevisiae*, there is a bias toward purines and uridines 3′ of the stop codon and a strong bias against the CG dinucleotide immediately following the terminator, which was proposed to cause faulty termination or failure of the ribosome to dissociate from the mRNA. Interestingly, it was noted that no yeast mRNA contains CGG or CGU immediately downstream of the stop codon for a major coding region and that uORF4 is unique in having the CGG triplet at this position. In contrast, the preference termination sequence derived in Table 3 that is permissive for reinitiation conforms to a typical termination region in *S. cerevisiae*. Perhaps the CGG triplet is avoided because it holds the ribosome in the termination region too long and this property is exploited by uORF4 to promote efficient ribosome dissociation from the mRNA. This explanation implies that translation of typical mRNAs can be stimulated by the retention of ribosomal subunits in the 3′ untranslated region following peptide chain release. Perhaps these subunits can be recycled efficiently to the 5′ end of the mRNA for new rounds of initiation. This proposal is in accord with the data in Table 4 indicating that the CGG triplet makes the largest contribution of the uORF4 termination region to the inhibition of reinitiation.

According to the models presented above, a sequence present in the termination region of uORF1 will be compatible with efficient reinitiation at *GCN4* so long as it does not interact with rRNA, mRNA, or RRF in a way that impedes the resumption of scanning following peptide release. The results shown in Table 2 suggest that a relatively large number of A+U-rich sequences, in addition to the termination sequence from wild-type uORF1, will fulfill these criteria. This might explain why replacing the CCG proline codon from uORF4 with the UAA or UGA stop codons restores efficient reinitiation in the presence of the termination sequence from uORF4 (Table 1). In addition to removing the CCG proline codon, these two substitutions change the sequence of the termination region, placing the previous stop codon (UAA) in the first three positions of the new termination region. The fact that the new termination region for these two constructs contains no G or C nucleotides in the critical block of three nucleotides immediately following the stop codon may account for their ability to restore reinitiation; alternatively, a two-codon uORF may be inherently more permissive than a three-codon uORF for reinitiation at *GCN4*.

The results of dissecting the termination sequence of uORF4 in Table 5 suggest that the number of termination sequences which can efficiently impede the resumption of scanning may be considerably smaller than that which is compatible with reinitiation, since two separate blocks of nucleotides in the uORF4 termination region contribute to its negative effect on reinitiation. In accord with this conclusion, using random mutagenesis, we were unable to identify single base substitutions in the uORF4 termination region that restored efficient reinitiation at *GCN4* (data not shown). However, the fact that the antipreference sequence derived in Table 3 does not show a strong resemblance to the uORF4 termination region may indicate that the ability of these two sequences to inhibit reinitiation involves distinctly different interactions with the translational machinery.

In summary, the propensity of the sequences immediately surrounding the terminator of uORF1 to promote efficient reinitiation at *GCN4* can be mimicked by numerous sequences whose only obvious defining feature is a high A+U

content. In fact, the UGC triplet normally present at uORF1 is much less permissive for reinitiation than a variety of more AU-rich triplets that we identified (Table 1). The sequences surrounding the termination codon of uORF4 are incompatible with reinitiation at *GCN4*, and it is now clear that the CCG codon in uORF4 makes a large contribution to this property. The CGG triplet immediately following the stop codon at uORF4 appears to be the most important element downstream from uORF4 that limits reinitiation. Replacing the C+G-rich sequences surrounding the uORF4 terminator with A+U-rich sequences restores high-level reinitiation at *GCN4*. On the basis of the strong correlation between high A+U content around the uORF stop codon and reinitiation downstream, we propose that reinitiation is impaired by stable interactions between nucleotides surrounding the stop codon with (i) the anticodon of the tRNA which decodes the third codon of the uORF, (ii) nucleotides in the rRNA, or (iii) nucleotides located elsewhere in *GCN4* mRNA. It is envisioned that these base-pairing interactions prevent the ribosome from exiting the termination region fast enough to escape the action of an RRF which would catalyze ribosome dissociation from the mRNA, thereby preventing reinitiation at *GCN4*. We now hope to use a genetic approach to identify individual components of the translational machinery which interact differently with the mRNA sequences surrounding the stop codons at uORF1 and uORF4 and are responsible for the distinctive abilities of these two uORFs to allow a resumption of scanning and reinitiation at downstream start codons.

ACKNOWLEDGMENTS

We thank Ronda Rolfes, Tom Dever, and Graham Pavitt for critical reading of the manuscript and Phil Farabaugh for the gift of plasmids.

REFERENCES

1. **Abastado, J. P., P. F. Miller, B. M. Jackson, and A. G. Hinnebusch.** 1991. Suppression of ribosomal reinitiation at upstream open reading frames in amino acid-starved cells forms the basis for *GCN4* translational control. Mol. Cell. Biol. **11:**486–496.
2. **Azad, A. A., and N. J. Deacon.** 1980. The 3′-terminal structure of five eukaryotic rRNAs determined by the direct chemical method of sequencing. The highly conserved sequences include an invarient region complementary to eukaryotic 5S rRNA. Nucleic Acids Res. **8:**4365–4376.
3. **Baim, S. B., and F. Sherman.** 1988. mRNA structures influencing translation in the yeast *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8:**1591–1601.
4. **Belcourt, M. F., and P. J. Farabaugh.** 1990. Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. Cell **62:**339–352.
5. **Brown, C. M., P. A. Stockwell, C. Trotman, and W. P. Tate.** 1990. Sequence analysis suggests that tetra-nucleotides signal the termination of protein synthesis in eukaryotes. Nucleic Acids Res. **18:**6339–6345.
6. **Brown, C. M., P. A. Stockwell, C. N. Trotman, and W. P. Tate.** 1990. The signal for the termination of protein synthesis in prokaryotes. Nucleic Acids Res. **18:**2079–2086.
7. **Cavenar, D. R., and S. C. Ray.** 1991. Eukaryotic start and stop translation starts. Nucleic Acids Res. **19:**3185–3192.
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. **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.
10. **Clemens, M. J., A. Galpine, S. A. Austin, R. Panniers, E. C. Henshaw, R. Duncan, J. W. Hershey, and J. W. Pollard.** 1987. Regulation of polypeptide chain initiation in Chinese hamster ovary cells with a temperature-sensitive leucyl-tRNA synthetase. Changes in phosphorylation of initiation factor eIF-2 and in the activity of the guanine nucleotide exchange factor GEF. J. Biol. Chem. **262:**767–771.
11. **Craigen, W. J., C. C. Lee, and C. T. Caskey.** 1990. Recent advances in peptide chain termination. Mol. Microbiol. **4:**861–865.
12. **Curran, J. F.** 1993. Analysis of effects of tRNA:message stability on frameshift frequency at the Escherichia coli RF2 programmed frameshift site. Nucleic Acids Res. **21:**1837–1843.
13. **Dever, T. E., L. Feng, R. C. Wek, A. M. Cigan, T. D. Donahue, and A. G. Hinnebusch.** 1992. Phosphorylation of initiation factor 2α by protein kinase GCN2 mediates gene-specific translational control of *GCN4* in yeast. Cell **68:**585–596.
14. **Folley, L. S., and M. Yarus.** 1989. Codon contexts from weakly expressed genes reduce expression in vivo. J. Mol. Biol. **209:** 359–378.
15. **Gafner, J., E. M. DeRobertis, and P. Phillippsen.** 1983. Delta sequences in the 5′ non-coding region of yeast tRNA genes. EMBO J. **2:**583–591.
16. **Guthrie, C., and J. Abelson.** 1982. Organization and expression of tRNA genes in *Saccharomyces cerevisiae*, p. 487–528. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), The molecular biology of the yeast *Saccharomyces*: metabolism and gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Hershey, J. W. B.** 1991. Translational control in mammalian cells. Annu. Rev. Biochem. **60:**717–755.
18. **Hill, J. R., and D. R. Morris.** 1993. Cell-specific translational regulation of S-adenosylmethionine decarboxylase mRNA. J. Biol. Chem. **268:**726–731.
19. **Hinnebusch, A. G.** 1984. Evidence for translational regulation of the activator of general amino acid control in yeast. Proc. Natl. Acad. Sci. USA **81:**6442–6446.
20. **Hinnebusch, A. G.** 1985. A hierarchy of *trans*-acting factors modulate translation of an activator of amino acid biosynthetic genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **5:**2349–2360.
21. **Hinnebusch, A. G.** 1992. General and pathway-specific regulatory mechanisms controlling the synthesis of amino acid biosynthetic enzymes in *Saccharomyces cerevisiae*, p. 319–414. *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.
22. **Hinnebusch, A. G., R. C. Wek, T. E. Dever, A. M. Cigan, L. Feng, and T. F. Donahue.** 1993. Regulation of *GCN4* expression in yeast: gene specific translational control by phosphorylation of eIF-2α, p. 87–115. *In* J. Ilan (ed.), Translational regulation of gene expression. Plenum Press, New York.
23. **Ichikawa, S., and A. Kaji.** 1989. Molecular cloning and expression of ribosome releasing factor. J. Biol. Chem. **264:**20054–20059.
24. **Ikemura, T.** 1982. Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes differences in synonymous codon choice patterns of yeast and *Escherichia coli* with reference to the abundance of isoaccepting transfer RNAs. J. Mol. Biol. **158:**573–597.
25. **Ito, H., Y. Fukada, K. Murata, and A. Kimura.** 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **153:**163–168.
26. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47:**1–45.
27. **Kozak, M.** 1986. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc. Natl. Acad. Sci. USA **83:**2850–2854.
28. **Kozak, M.** 1987. Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. Mol. Cell. Biol. **7:**3438–3445.
29. **London, I. M., D. H. Levin, R. L. Matts, N. S. B. Thomas, R. Petryshyn, and J. J. Chen.** 1987. Regulation of protein synthesis,

p. 359–380. *In* P. D. Boyer and E. G. Krebs (ed.), The enzymes, vol. 18. Academic Press, New York.

30. **Miller, P. F., and A. G. Hinnebusch.** 1989. Sequences that surround the stop codons of upstream open reading frames in *GCN4* mRNA determine their distinct functions in translational control. Genes Dev. **3:**1217–1225.
31. **Mueller, P. P., S. Harashima, and A. G. Hinnebusch.** 1987. A segment of *GCN4* mRNA containing the upstream AUG codons confers translational control upon a heterologous yeast transcript. Proc. Natl. Acad. Sci. USA **84:**2863–2867.
32. **Mueller, P. P., and A. G. Hinnebusch.** 1986. Multiple upstream AUG codons mediate translational control of *GCN4*. Cell **45:** 201–207.
33. **Mueller, P. P., B. M. Jackson, P. F. Miller, and A. G. Hinnebusch.** 1988. The first and fourth upstream open reading frames in *GCN4* mRNA have similar initiation efficiencies but respond differently in translational control to changes in length and sequence. Mol. Cell. Biol. **8:**5439–5447.
34. **Murgola, E. J., K. A. Hijazi, H. U. Goringer, and A. E. Dahlberg.** 1988. Mutant 16S ribosomal RNA: a codon specific translational suppressor. Proc. Natl. Acad. Sci. USA **85:**4162–4165.
35. **Pain, V. M.** 1986. Initiation of protein synthesis in mammalian cells. Biochem. J. **235:**625–637.
36. **Parent, S. A., C. M. Fenimore, and K. A. Bostian.** 1985. Vector systems for the expression, analysis and cloning of DNA sequences in *S. cerevisiae*. Yeast **1:**83–138.
37. **Pelletier, J., and N. Sonenberg.** 1985. Insertion mutagenesis to increase secondary structure within the 5′ noncoding region of a eukaryotic mRNA reduces translational efficiency. Cell **40:**515–526.
38. **Persson, B. C.** 1993. Modification of tRNA as a regulatory device. Mol. Microbiol. **8:**1011–1016.
39. **Ryoji, M., J. W. Karpen, and A. Kaji.** 1981. Further characterization of ribosome releasing factor and evidence that it prevents ribosomes from reading though a termination codon. J. Biol. Chem. **256:**5798–5801.
40. **Sanger, F. S., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
41. **Schleiss, M. R., C. R. Degnin, and A. P. Geballe.** 1991. Translational control of human cytomegalovirus gp48 expression. J. Virol. **65:**6782–6789.
42. **Sharp, P. M., and W.-H. Li.** 1987. The codon adaptation index-a measure of directional synonomous codon usage, and its potential applications. Nucleic Acids Res. **15:**1281–1295.
43. **Sharp, P. M., T. M. F. Tuohy, and K. R. Mosurski.** 1986. Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. Nucleic Acids Res. **14:**5125–5143.
44. **Turner, D. H., and N. Sugimoto.** 1988. RNA structure prediction. Annu. Rev. Biophys. Biophys. Chem. **17:**167–192.
45. **Valle, R. P., and M.-D. Morch.** 1988. Stop making sense or regulation at the level of termination in eukaryotic protein synthesis. FEBS Lett. **235:**1–15.
46. **Weiss, G. B.** 1973. Translational control of protein synthesis by tRNA unrelated to changes in tRNA concentration. J. Mol. Evol. **2:**199–204.
47. **Werner, M., A. Feller, F. Messenguy, and A. Pierard.** 1987. The leader peptide of yeast gene *CPA1* is essential for the translational repression of its expression. Cell **49:**805–813.
48. **Wolin, S. L., and P. Walter.** 1988. Ribosome pausing and stacking during translation of a eukaryotic mRNA. EMBO J. **7:**3559–3569.
49. **Xu, H., and J. D. Boeke.** 1990. Host genes that regulate transposition in yeast: the abundance of a rare tRNA regulates Ty1 transposition frequency. Proc. Natl. Acad. Sci. USA **87:** 8360–8364.
50. **Yoon, H., and T. F. Donahue.** 1992. Control of translation initiation in *Saccharomyces cerevisiae*. Mol. Microbiol. **6:**1413–1419.