Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 4849–4853, July 1986
Genetics

Multiple *cis*-acting elements modulate the translational efficiency of GCN4 mRNA in yeast

(deletion-insertion mutation/*lacZ* fusion/*GCN2, GCD1* genes/upstream AUG codon)

DIMITRIS TZAMARIAS*†, DESPINA ALEXANDRAKI*, AND GEORGE THIREOS*†

*Institute of Molecular Biology and Biotechnology, Research Center of Crete, P.O. Box 1527, Heraklio, Crete, Greece; and †Department of Biology, University of Crete, Heraklio, Crete, Greece

Communicated by Fotis C. Kafatos, March 10, 1986

ABSTRACT **The expression of the *GCN4* gene in yeast is regulated at the translational level: growth of yeast cells under amino acid starvation conditions results in an increase in the translational efficiency of GCN4 mRNA. A sequence within the 5′ untranslated region of this mRNA, which contains four small open reading frames, acts in *cis* to suppress translation when growth occurs in rich media. In this report, we have analyzed the effects on translation of a series of deletion, insertion, and substitution mutations in the 5′ untranslated region of GCN4 mRNA. This analysis showed that at least two distinct *cis* elements located within the region of the small upstream open reading frames are required, in conjunction with *trans* positive elements, for the translational activation of GCN4 mRNA. We propose that the translational efficiency of GCN4 mRNA is modulated by the rate of translation initiation at the upstream AUG codons.**

The *GCN4* gene of *Saccharomyces cerevisiae* encodes a protein required for transcriptional activation of a number of genes involved in amino acid biosynthesis (1, 2). Such activation possibly results from the specific binding of *GCN4*-encoded protein to the promoter region of co-regulated genes (3). The expression of *GCN4* gene is subject to translational regulation (4, 5): when yeast cells are grown under amino acid starvation conditions the translational efficiency of GCN4 mRNA is increased severalfold. The translational repression of this mRNA in rich medium is attributable at least in part to a sequence present within its long 5′ untranslated region, which contains four small open reading frames (ORFs) (4, 5). Translational derepression of the mRNA requires the function of at least three unlinked positive regulatory genes, *GCN1, GCN2*, and *GCN3*. Mutations in a fourth regulatory gene, *GCD1*, result in constitutively derepressed translation of GCN4 mRNA (6). This system is one of the few clear examples of gene regulation at the translational level in eukaryotes, and it offers an opportunity for elucidation of the mechanisms involved (7). In this report, we analyze *in vivo* the effects of deletion, insertion, and substitution mutations in the 5′ untranslated region of GCN4 mRNA, and thereby we identify two *cis* elements that are required for its translational derepression.

MATERIALS AND METHODS

Strains. The *gcd1* strain was obtained from Kevin Struhl and has the genotype *a gcd1-1 ura3-52 can1*. All other yeast and bacterial strains used have been described previously (2).

Mutant Constructions. The Δ1 mutation of the *GCN4* gene lacks sequences from position +166 to position +445 of the 577-nucleotide-long 5′ untranslated region (4). Substitution mutations were generated by insertion of DNA sequences into the *Sst* I site of the Δ1 mutation. (The *Sst* I site was generated close to the destroyed *Bst*EII site as a result of the Δ1 construction.) A series of deletion mutations was generated using the deletion subclones E9, E7, and E5, which were constructed for sequencing purposes (4). One endpoint of these BAL-31-generated deletions is within the 5′ untranslated region of the *GCN4* gene, and the other endpoint is a *Hin*dIII site in the 5′ flanking region of the *GCN4* gene (position −552) (4). Fragments bearing these deletions were inserted into the *Hin*dIII–*Sma* I sites of plasmid pUC18. The *Hin*dIII–*Eco*RI fragment (the *Eco*RI site was derived from the vector) of each one of these subclones was used to replace the *Hin*dIII–*Bst*EII fragment of the wild-type gene and produced deletion mutants Δ1, Δ5, and Δ9. Deletion mutants Δ9S and Δ7S were constructed by inserting the *Sau*3AI–*Bst*EII fragment of *GCN4* into the *Sst* I site of subclones E9 and E7 and using the *Hin*dIII–*Eco*RI fragments of the resulting subclones to replace the *Hin*dIII–*Bst*EII fragment of the wild-type gene. The Δ13 mutation was generated as follows: the subcloned *Sau*3AI–*Bst*EII fragment at the *Kpn* I site of pUC18 was used to generate deletion mutations from the *Sau*3AI end by BAL-31 digestion, the resulting deletion mutations were excised by *Kpn* I digestion of their other end, and the fragments obtained were introduced into the *Sst* I site of the E7 subclone. Thus, the Δ13 mutation was generated by replacing the *Hin*dIII–*Bst*EII fragment of the wild-type gene by the *Hin*dIII–*Eco*RI fragment of one of the above resulting subclones. The right endpoint of the Δ13 mutant was determined by DNA sequencing. All of the above described deletion mutations contain a small sequence derived from the pUC18 polylinker between the *Sma* I and the *Eco*RI sites (15 bases) that did not introduce any AUG codons. Deletion mutants Δ33 and Δ35 were generated by digestion with BAL-31 in both directions, starting at the *Bst*EII site, and rejoining the digested molecules. The endpoints of these deletions were determined by DNA sequencing. Insertion mutations were generated by introducing various DNA sequences at the *Bst*EII site of the *GCN4* gene.

Plasmid Constructions. All mutant genes were fused in-frame with the *lacZ* gene of *Escherichia coli*, as described for the wild-type *GCN4* gene (4). In these fusions, the 5′ flanking region of the *GCN4* gene extending to nucleotide −552 (the *Hin*dIII site), the 5′ untranslated region resulting from each mutation, and the portion of the *GCN4* coding region encoding the 53 amino-terminal amino acids were fused in-frame with the *lacZ* gene. The fused sequences were introduced into the single copy yeast vector YCp50, which also contained the yeast *URA3* gene as a selectable marker.

Translation Assays. The *lacZ* fusion-containing plasmids were used to transform a *ura3-52* (wt), a *gcn2 ura3-52 (gcn2)*, and a *gcd1 ura3-52 (gcd1)* strain by standard procedures (8). Transformed yeast cells were grown in minimal medium containing all 20 amino acids, and β-galactosidase activity

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: ORF, open reading frame.

was assayed as described by Miller (9). Amino acid starvation conditions were accomplished by transferring yeast cells grown under repressing conditions into minimal medium containing 10 mM 3-amino-1,2,4-triazole, which elicits histidine starvation (10). β-Galactosidase activity was assayed after 4 hr of growth under such derepressing conditions.

RESULTS

Translational Efficiency Assays. In our analysis, translational efficiencies were calculated by measuring *in vivo* levels of β-galactosidase activity after in-frame fusion of wild-type *GCN4* sequences or mutant derivatives to the coding region of the *E. coli lacZ* gene followed by introduction into yeast cells. The normal sequence of the 5′ untranslated region of GCN4 mRNA is shown in Fig. 1*a*. The *GCN4-lacZ* constructs were introduced into both wild-type and mutant yeast strains. In a *gcn2* strain, the wild-type fusion mRNA is translated at low basal levels under repressing conditions and no increase in translational efficiency occurs under amino acid starvation conditions. Thus the *gcn2* background is characterized by the absence of a positive regulator of GCN4 mRNA translation (6). In a *gcd1* strain, the wild-type fusion mRNA exhibits constitutively high levels of translation (ref. 6 and unpublished results). Thus, the *gcd1* background is characterized by the absence of a negative regulator. This regulator either has a direct effect on the translation of GCN4 mRNA or it has an indirect effect on translation by suppressing the function of a positive regulator(s). The latter alternative is supported by our results.

β-Galactosidase activity was measured in transformed cells grown in medium containing all 20 amino acids (repressing conditions) or in medium eliciting histidine starvation (derepressing conditions). To limit our attention to translational regulation, the levels of β-galactosidase activity were corrected for the relative increase in fusion mRNA levels under derepressing conditions and for the variability of these levels in the different genetic backgrounds, as determined by DNA·RNA hybridization analysis (data not shown). There was no observable variation of those levels among the various mutant GCN4-lacZ mRNAs; the strain and growth condition variations were very close to those reported elsewhere (6).

Substitution Mutations. It has previously been reported (4) that deletion mutant Δ1 (see Figs. 1*a* and 2*b*), which lacks all four small ORFs present in the 5′ untranslated region of the GCN4-lacZ mRNA, is translated at constitutively high levels in all strain backgrounds. In the first series of mutations reported here, we substituted the region containing the four ORFs by inserting into deletion mutant Δ1 three different sequences (S1, S2, and S3 mutants). S1 contains a 115-base-pair *Alu* I fragment derived from the 3′ end of a sea urchin α-tubulin cDNA clone (11) (from codon 427, extending 38 nucleotides into the 3′ untranslated sequence) and including an open reading frame 8 codons long. The substituted sequence of S1 is shown in Fig. 1*b*. S2 contains the S1 substitution sequence inserted in the reverse orientation, whereas S3 contains the sequence removed in the Δ1 deletion inserted in the reverse orientation. The S3 substitution sequence has two small open reading frames, one consisting of 12 codons and a second one initiating at the end of this inverted sequence and terminating further downstream after 8 codons. The substitution mutations are shown schematically in Fig. 2*a*. Substitutions using the S1 and S3 sequences, containing one and two ORFs, respectively, drastically reduced the translational efficiency of the fusion mRNAs in all three backgrounds (Table 1). By contrast, the introduction

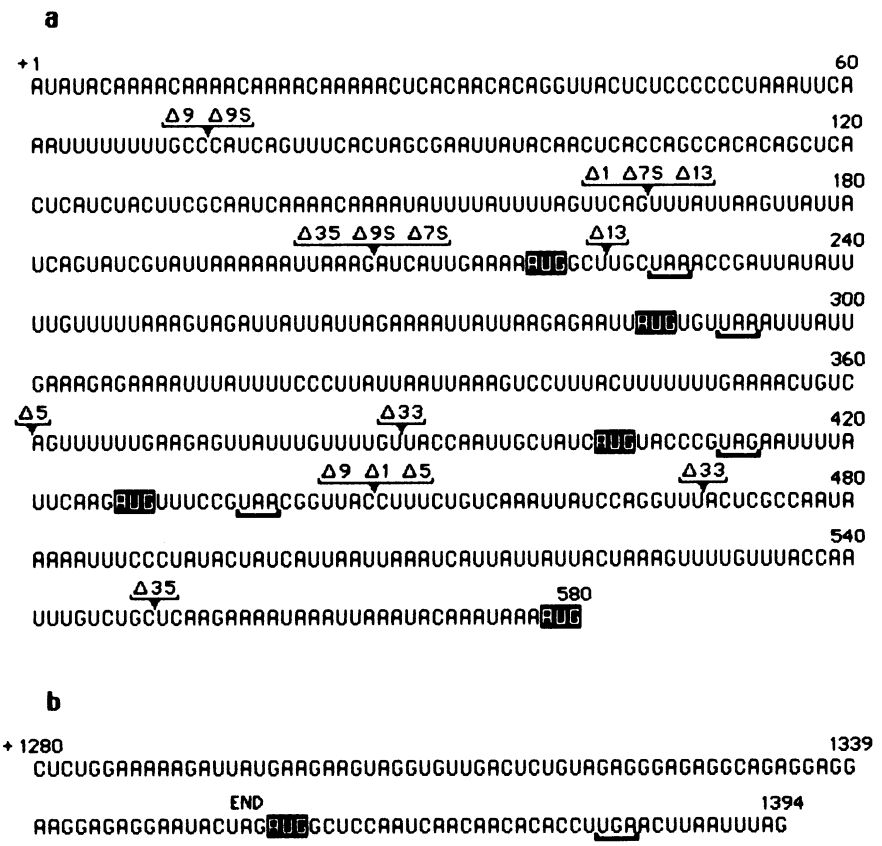

FIG. 1. (*a*) Nucleotide sequence of the 5′ untranslated region of GCN4 mRNA (data derived from ref. 3). The positions of the end points for each deletion constructed are indicated. Initiation codons are indicated by black boxes, whereas termination codons are underlined. The AUG at position 578 is the start codon of the *GCN4* gene. (*b*) Nucleotide sequence of the α-tubulin cDNA fragment used in the substitution and insertion mutations (10). Initiation and termination codons are shown as in *a*.

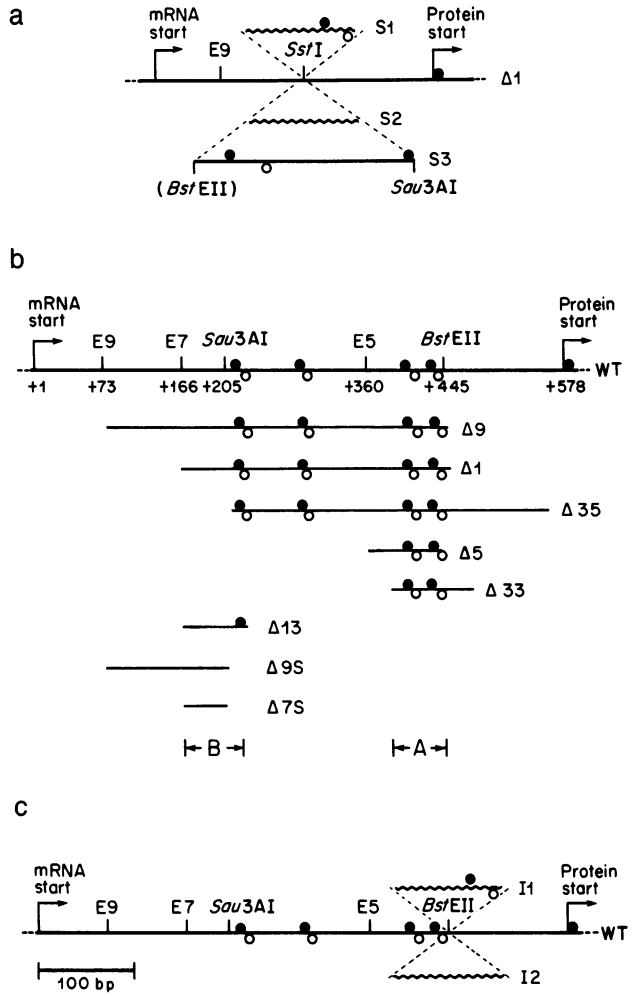

FIG. 2. Schematic representation of the mutations constructed in the 5′ untranslated region of the *GCN4* gene. (*a*) Substitution mutations. The 5′ untranslated region in the Δ1 deletion mutant of the *GCN4* gene is shown. S1, S2, and S3 DNA fragments were inserted at the *Sst* I site of the Δ1 mutant *GCN4* gene. (*b*) Deletion mutations. A map of the 5′ untranslated region of the wild-type (WT) *GCN4* gene is presented. The numbers indicate the sequence positions of the sites shown. E9, E7, and E5, end points of deletions generated for sequencing purposes (3). The extent of each of the deletions is shown below the map. Area A is defined by the common overlapping sequence of mutants Δ5 and Δ33, whereas area B is defined by the Δ13 deletion. (*c*) Insertion mutations. These mutants were generated by inserting a DNA fragment (S1) used in the substitution mutations into the *Bst*EII site of the *GCN4* gene in both orientations (I1 and I2). ●, AUG codons; ○, termination codons.

of the S2 substitution sequence, which does not contain any ORF, resulted in β-galactosidase levels similar to those of the Δ1 mutant, although somewhat lower in all backgrounds, probably because of a sequence-specific or length-specific effect or both. In agreement with evidence from other eukaryotic systems (12, 13), these results show that the introduction of small ORFs in the 5′ untranslated region leads to translational repression. Furthermore, these results support the hypothesis that ORFs account for the negative effect of the sequence defined by the Δ1 deletion on the translation of GCN4 mRNA. Unlike wild-type GCN4 mRNA, neither S1 or S3 substituted mutant mRNAs could be translationally derepressed, showing that the negative effect of foreign ORFs on translation cannot be alleviated by amino acid starvation. Thus translational derepression must depend on *cis* sequences present in the Δ1 deleted region.

Deletion Mutations. To define sequences within the Δ1 deleted region that are responsible for the positive regulation of translation, a series of deletion mutations was constructed (Fig. 2*b*) and the effects of these mutations on translational efficiency were assayed (Table 1). In agreement with our previous conclusions, any deletion that removed sequences including all four ORFs (Δ9, Δ1, and Δ35) resulted in an elevated constitutive translational efficiency, bypassing the necessity for the *GCN2* function. Deletions of sequences in area A, which included the two ORFs most proximal to the protein start (Δ5, Δ33), resulted in elevated basal translational levels in both the wild-type and the *gcn2* backgrounds (Table 1). We interpret this to be the direct consequence of removing the negative effects on translation of the two ORFs included in these deletions.

Furthermore, in such mutants translational derepression did not occur in either the wild-type or the *gcn2* strains, suggesting that elements within the deleted sequences are required for translational regulation. The phenotype of the Δ5 deletion persisted even when the deleted region was replaced by the S2 substitution sequence, which is roughly the same size, suggesting that this phenotype is not a consequence of a position effect (data not shown). Finally, the translational efficiency of both the Δ5 and Δ33 mutants in the *gcd1* strain was decreased by a factor of almost 5, suggesting that the *gcd1* background is characterized by the absence of a negative regulator that results in the constitutive presence of one or more positive regulators of translation that require area A for their function.

Deletion of sequences starting from the E7 site and ending within the 5′ most proximal ORF (Δ13), which covers area B (Fig. 2*b*), resulted in a dramatic reduction of the basal levels of translation, even though the number of ORFs was decreased. The extent of translational derepression of this mutant mRNA was also affected: the translation of Δ13 increased by a factor of 4, whereas the translation of wild-type increased by a factor of 10. Deletions in the same region ending 11 bases from the first AUG (Δ7S, Δ9S) affected the basal levels to a smaller degree but regulation to the same extent as seen in the Δ13 mutation (Table 1). These results suggest that an element within the 17-base-pair sequence defined by the *Sau*3AI site and the right end point of the Δ13 deletion has a strong positive effect on basal levels, whereas a sequence in the overlapping region extending to position E7 is required to establish the wild-type extent of translational derepression. Finally, area B has a very strong effect in the *gcd1* background; its deletion reduces the translational efficiency to very low levels despite the presence of area A. Thus, sequences in region B are required for *GCD1*-mediated control. The results of deletions in area B suggest that positive regulation at this region has a significant effect on the establishment of wild-type basal levels and the final derepressed levels of translation.

Insertion Mutations. The third type of mutation that we analyzed was the insertion of foreign DNA sequences into the region between the 3′ most proximal ORF and the point of protein start (Fig. 2*c*). Two insertion mutants were analyzed, one containing and one lacking an additional small ORF. Both resulted in reduction of both the basal and the derepressed levels of translation. The insertion mutants were poorly translated in the *gcd1* background. In agreement with our previous observations, the ORF-containing insertion had a more pronounced effect. It seems that inserted sequences may directly or indirectly interfere with the function of the upstream regulatory elements.

DISCUSSION

At least three elements within the 5′ untranslated region of GCN4 mRNA are required in *cis* for its translational modulation. One element is defined by the Δ1 deleted sequence, which determines the low basal levels of translation under

Table 1. Normalized levels of β-galactosidase activity produced by wild-type and mutant *GCN4-lacZ* fusions in different strain backgrounds and under different growth conditions

Strain/ fusion	Description	β-Galactosidase activity, relative units								
		Wild type			*gcn2*			*gcd1*		
		R	D	D/R	R	D	D/R	R	D	D/R
wt	Four normal ORFs	1.0	10.4	10.4	0.2	0.2	1.0	185	171	0.9
Δ1	No ORFs	75	71	0.9	79	46	0.6	167	168	1.0
Δ9	Extended Δ1; no ORFs	65	54	0.8	74	49	0.7	163	175	1.1
Δ35	Extended Δ1; no ORFs	66	71	1.1	65	41	0.6	83	92	1.1
S2	Tubulin substitution; no ORFs	54	46	0.8	48	33	0.7	58	68	1.2
S1	Tubulin substitution; one ORF	0.7	0.7	1.0	0.6	0.6	0.9	0.7	0.8	1.1
S3	Inversion; two ORFs	1.2	1.6	1.3	2.0	1.0	0.5	0.7	0.8	1.2
Δ5	Deletion, area A; two ORFs	7.5	9.2	1.2	9	5.4	0.6	33	38	1.1
Δ33	Deletion, area A; two ORFs	8.3	8.8	1.1	10	5.3	0.5	43	41	1.0
Δ13	Deletion, area B; three ORFs	0.2	0.7	4.0	0.1	0.2	1.7	0.5	0.4	0.8
Δ7S	Deletion, area B; four ORFs	0.7	2.3	3.4	0.3	0.2	0.5	0.6	0.6	1.0
Δ9S	Deletion, area B; four ORFs	0.7	2.4	3.6	0.3	0.3	0.8	0.6	0.5	0.9
I1	Tubulin insertion; five ORFs	0.2	0.4	2.1	0.1	0.2	1.6	0.2	0.3	1.5
I2	Tubulin insertion; four ORFs	0.4	1.3	3.0	0.1	0.1	1.2	1.5	1.8	1.2
URA3	Control	1.6	1.5	0.9	1.6	1.3	0.8	1.7	1.8	1.1

β-Galactosidase activity was measured after growth of transformed yeast cells in medium containing all 20 amino acids (R) or in medium that elicits histidine starvation (D). Levels of β-galactosidase activity were normalized for levels of each fusion mRNA in each genetic background and under each growth condition to basal levels of the wild-type GCN4-lacZ mRNA in the wild-type strain. Actual units of β-galactosidase activity were divided by a factor of 1.2 to express the wild-type fusion basal levels as 1. D/R ratios were calculated prior to rounding off numbers. The results using a *URA3-lacZ* fusion gene in YCp50 are included as a control. The lower translational levels in the *gcn2* background under derepressing conditions are the consequence of poor growth of this strain under amino acid starvation conditions.

repressing conditions. Our results strongly suggest that the negative effect of this element is a consequence of the existence of the four ORFs within it. Substitution of this region with other foreign sequences containing ORFs resulted in basal levels similar to wild type, while substitution with a sequence lacking ORFs resulted in high translation levels. Naturally, the foreign ORF-containing sequences did not restore translational derepression in a wild-type background and did not allow constitutive derepression of translation in the *gcd1* background. These data suggest that specific elements within the Δ1 deleted region are required for positive translational regulation of GCN4 mRNA.

Indeed, our analysis of the effects of mutations in the 5′ untranslated region of GCN4 mRNA identified two elements acting in *cis*, which are required for the translational derepression of this mRNA. One element, within area A, may define a site of interaction with *trans* activators of translation. In addition, area A has a negative effect on the basal levels of translation, accounted for by the presence of two small ORFs within it. A second element, within area B, is required for the establishment of appropriate basal levels of translation and partially affects the extent of translational derepression. The function of both elements is inhibited by downstream insertions.

Our results suggest that the activity of these *cis* elements requires the function of *trans* positive regulators. Deletion of area A resulted in reduction by a factor of 5 of the translational levels in the *gcd1* strain background, while deletion of area B drastically suppressed translation of the fusion mRNA in this background. These results show that a nonfunctional *GCD1* gene product results in the activation of one or more *trans* activators of translation. Furthermore, the activator(s) that acts on area B is a prerequisite for the highly derepressed translational levels in the *gcd1* background. In a *gcn2 gcd1* double mutant strain the levels of translation of the wild-type fusion mRNA are lower than those in the *gcd1* strain by a factor of 3 to 4 (6). These levels are similar to those observed in the *gcd1* background with the fusion mRNAs carrying deletions in area A. These observations suggest that the GCN2-dependent activator acts primarily in area A. This is supported by our results showing that the absence of *GCN2* enhances the effect of deletions in area B on basal levels of translation, while there is no additional effect in the absence of *GCN2* on the levels resulting from deletions in area A. Thus there appear to be at least two distinct positive regulators of translation, one acting on area B to establish primarily the basal levels of translation of *GCN4* mRNA (*GCD1* dependent) and the other acting on area A in response to amino acid starvation (*GCN2* dependent).

The modified ribosome scanning hypothesis for translation initiation in eukaryotes (13) postulates that the rate of initiation of translation in a eukaryotic mRNA depends on the stability of attachment of 40S ribosomal subunits migrating down the 5′ untranslated region. Based on this hypothesis, we propose that the basal levels of GCN4 mRNA translation are determined in part by the interaction of a *trans* regulator with area B, resulting in the stabilization of 40S subunits. Under repressing conditions the association of 40S subunits with the 5′ untranslated region is destabilized because of the successive rounds of initiation–termination further downstream, and thus the final number of subunits reaching the "real" AUG is diminished. The Δ13 deletion extends only 17 bases beyond the *Sau*3AI site, and yet it has a significant effect on the basal translation levels, compared to the Δ7S deletion, indicating that this stabilization could be indirectly determined by the rate of translation initiation at the first "false" AUG. It might be relevant that this 5′ most proximal AUG has the optimum sequence context for translation initiation (14) whereas the other four AUGs have suboptimal sequence contexts. In this respect, the results with deletion Δ7S could be attributed to a change in the flanking sequence context of this AUG. Under derepressing conditions, there is an increased stabilization of the 40S subunits, primarily in area A and partially in area B, accomplished by the increased activity of *trans* regulators interacting with these regions, resulting in an increase in the rate of initiation at the real AUG. An intriguing and testable possibility is that derepression is accomplished by an increase in the rate of translation initiation at the upstream false AUGs, which could increase the number of 40S subunits remaining attached. In this

respect, the phenotype of the insertion mutations can be explained as a result of destabilization of the 40S subunit because of the additional "foreign" sequence that was inserted.

Our analysis is not exhaustive, and additional *cis* regulatory elements might be required. Preliminary results indicate that sequences within the region defined by the *Bst*EII site and the protein start also have an effect on both basal levels and regulation, which could explain the lower translational levels of the Δ35 deletion mutant in the *gcd1* background as compared to the levels of the Δ1 and Δ9 deletion mutants in the same strain. Additional mutation analysis, including point mutagenesis within the identified elements, will provide more information on the mechanisms that regulate expression of the *GCN4* gene at the translational level.

We thank F. C. Kafatos for helpful suggestions, Kevin Struhl and Mark Rose for strains and plasmids, Bernie Galgoci for artwork, and Martha Peterson for typing the manuscript. This work has been supported by funds provided by the Greek Ministry of Industry, Research, and Technology.

1. Hinnebusch, A. G. & Fink, G. R. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5374–5378.
2. Penn, M. D., Thireos, G. & Greer, H. (1984) *Mol. Cell. Biol.* **4**, 520–528.
3. Hope, I. A. & Struhl, K. (1985) *Cell* **43**, 177–188.
4. Thireos, G., Penn, M. D. & Greer, H. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5096–5100.
5. Hinnebusch, A. G. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6442–6446.
6. Hinnebusch, A. G. (1985) *Mol. Cell. Biol.* **5**, 2349–2360.
7. Hunt, T. (1985) *Nature (London)* **316**, 580–581.
8. Hinnen, A., Hicks, J. B. & Fink, G. R. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1929–1933.
9. Miller, J. H. (1972) in *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 352–357.
10. Klopotowski, T. & Wiater, A. (1965) *Arch. Biochem. Biophys.* **112**, 562–566.
11. Alexandraki, D. & Ruderman, J. V. (1983) *J. Mol. Evol.* **19**, 397–410.
12. Liu, C., Simousen, C. C. & Levinson, A. D. (1984) *Nature (London)* **309**, 82–85.
13. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 3873–3893.
14. Kozak, M. (1983) *Microbiol. Rev.* **47**, 1–45.