MOLECULAR AND CELLULAR BIOLOGY, Jan. 1987, p. 104–110
0270-7306/87/010104-07$02.00/0

Vol. 7, No. 1

Two Related Regulatory Sequences Are Required for Maximal Induction of *Saccharomyces cerevisiae his3* Transcription

KEVIN STRUHL* AND DAVID E. HILL†

Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02115

Received 14 July 1986/Accepted 16 September 1986

In *Saccharomyces cerevisiae*, the coordinate induction of *his3* and other amino acid biosynthesis genes is mediated by the binding of GCN4 activator protein to specific promoter sequences. The *his3* regulatory region contains the sequence TGACTC, which with some variation is repeated six times upstream of the mRNA initiation site. The requirements for maximal *his3* induction were examined with a series of sequential 5′ deletion mutations as well as a set of small internal deletions. Deletions encroaching as far downstream as position −142 behave indistinguishably from the wild-type gene, thus indicating that the two proximal copies of the regulatory sequence are sufficient for maximal induction. Deletions with breakpoints between −137 and −99 confer inducibility, but not to the normal wild-type level. A deletion ending immediately upstream of the proximal TGACTC sequence (position −99) shows some constitutive expression that is independent of the *gcn4* gene product. Deletions extending to −94 or beyond do not produce detectable levels of *his3* mRNA. Small internal deletions that only remove the proximal regulatory sequence and a 1-base-pair deletion of the thymine residue at −99 abolish induction, but do not affect the basal level of transcription. These results indicate that the proximal copy between −99 and −94 is absolutely required for *his3* induction, whereas the copy between −142 and −137 is required only for the maximal level of induction and is inactive by itself. From these and other observations, we suggest the possibility that these related regulatory sequences may be targets for two distinct proteins.

Under normal growth conditions, the *Saccharomyces cerevisiae his3* gene is transcribed with equal efficiency from two distinct initiation sites (defined as +1 and +12) at a basal level of one or two mRNA molecules per cell (18, 20). Extensive deletion analysis indicates that constitutive *his3* transcription depends on a poly(dA-dT) upstream element located between −113 and −129 as well as a TATA element located between −35 and −55 (14, 17).

When yeast cells are starved for any one of several different amino acids, they respond by coordinately inducing the transcription of many amino acid biosynthesis genes scattered throughout the genome (reviewed in reference 10). Under these circumstances, *his3* transcription is induced by an overall factor of 3 above the basal level (20). Surprisingly, the *his3* transcript initiating at +12 (and a minor transcript initiating at +22) is induced by a factor of 5, whereas the +1 transcript is not induced at all (19).

By analysis of a set of *his3* deletion mutations for their ability to induce transcription in response to amino acid starvation, a *cis*-acting regulatory region was located between −83 and −103 (15). This region contains the sequence TGACTC, which is present in the promoter regions of essentially all genes regulated by amino acid starvation (1, 4a, 6, 15), and which is also critical for the induction of *his4* transcription (1, 8). The TGACTC regulatory sequence is a target for the specific binding by the GCN4 activator protein (9).

It has been noted that the TGACTC sequence, with minor variations, is often repeated at a promoter of a given coregulated gene (1, 6, 15). Analysis of *his4* derivatives suggested that there are three copies that serve as functional regulatory sites and that multiple copies are necessary for maximal induction (8). However, the interpretation of these experiments is complicated for two reasons. First, the *his4* deletions were not controlled for the distance between the different TGACTC sequences or their distance to the TATA element. Second, many of the derivatives strongly affected the basal level of transcription and hence produced variable induction ratios. As the elements necessary for basal level *his4* transcription were not defined, it was sometimes difficult to interpret the relationship between effects on basal or inducible *his4* transcription.

The *his3* promoter region contains two TGACTC sequences (positions −99 to −94 and −258 and −263) as well as four other sites with only a single deviation (−142 to −137, −181 to −176, −216 to −221, and −225 to −230) (18) (Fig. 1). Although some of these copies are oriented in the opposite direction with respect to the mRNA initiation sites, it appears that the TGACTC sequence can function in either orientation (8).

The observation that a deletion removing the region between −83 and −103 abolishes *his3* induction suggests that the proximal copy is critical and that the remaining five copies are irrelevant (15). This interpretation is strongly supported by the fact that in vitro GCN4 protein binds with an affinity of 10^{-10} M to the proximal copy, whereas it fails to bind detectably to any of the other copies (9). Thus, one might predict that the proximal TGACTC sequence and the TATA element would be sufficient for maximal *his3* induction.

In this study, we analyzed a series of deletion mutations that successively remove sequences upstream of the *his3* TATA element for their ability to induce *his3* transcription. In addition, we examined small deletions of the proximal TGACTC sequence including a 1-base-pair deletion of the upstream-most thymine residue. The results indicate that

* Corresponding author.
† Present address: Genetics Institute, Cambridge, MA 02140.

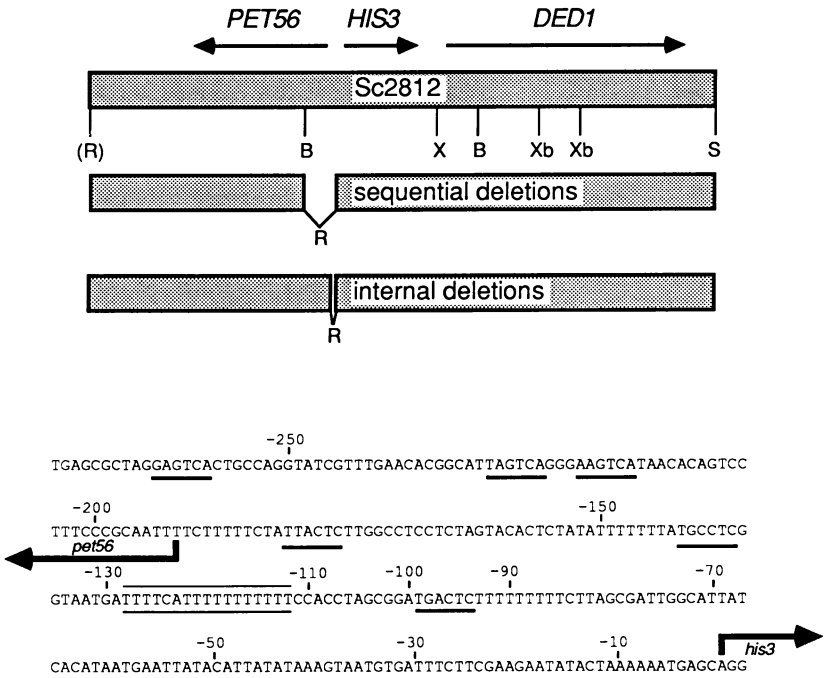

FIG. 1. DNA structures. The shaded bar at the top represents a 6.1-kbp *Eco*RI-*Sal*I fragment (with a mutated *Eco*RI site) that contains the intact *pet56*, *his3*, and *ded1* genes (18). The location and orientation of the transcripts are indicated by arrows above the shaded bar, and restriction endonuclease cleavage sites are indicated as vertical lines below the bar (R, *Eco*RI; B, *Bam*HI; X, *Xho*I; Xb, *Xba*I; S, *Sal*I). The shaded bars beneath Sc2812 represent the structures of the sequential and internal deletion mutations. The bottom half of the figure shows the nucleotide sequence of the *his3* coding strand of the promoter region (18). The coordinates above the sequence are determined with respect to the upstream-most *his3* initiation site which is defined as +1. Features of the DNA sequence include the *his3* and *pet56* initiation sites (thick arrows) (18), the poly(dA-dT) upstream element for constitutive expression (thin lines above and below the sequence) (17), and the six TGACTC and related sequences (thick lines below the sequence). The three most distal copies of this sequence are oriented in the opposite direction with respect to the mRNA initiation sites.

although the proximal copy is essential for *his3* induction, it is not sufficient for achieving the maximal wild-type level.

MATERIALS AND METHODS

Yeast strains. The yeast strains used in these experiments were KY117 (**a** *ura3-52 ade2-101 lys2-801 trp1-Δ1 his3-Δ200*) (17), KY114 (a *HIS3* derivative of KY117) (17), KY466 (**a** *ura3-52 lys2-801 his3-Δ200 gcd1-1*), and KY498 (**a** *ura3-52 trp1-Δ1 lys2-801 his3-Δ200 gcn4-2*). The *his3-Δ200* mutation deletes the entire *his3* structural gene and promoter region (17). Generally, these strains were grown at 30°C in liquid suspension or on 2% Bacto-Agar plates (Difco Laboratories, Detroit, Mich.) in YPD broth or in minimal medium containing 2% glucose and appropriate supplements (16). Strains containing the *gcd1-1* mutation are temperature sensitive for growth (22); hence, they were propagated at 23°C.

Construction of hybrid DNA molecules. The procedures for constructing hybrid DNA molecules have been described previously (16–18). All the molecules contain a 6.1-kilobase-pair (kbp) segment of yeast chromosomal DNA with the entire *pet56-his3-ded1* gene region cloned in the *ura3*$^+$ vectors YIp5 or YIp55 (18) (Fig. 1). The nucleotide sequences of the *his3* deletion alleles are shown in Fig. 2; they were created as follows.

For the set of sequential 5′ deletions, the 1.8-kbp *Bam*HI fragment containing the entire *his3* gene, Sc2676 (18), was treated for various amounts of time with *Bal* 31 nuclease. The resulting collection of fragments was treated with DNA polymerase I to blunt the ends, ligated to an octanucleotide *Eco*RI linker (GGAATTCC), cleaved with *Eco*RI and *Xho*I, and ligated between the *Eco*RI and *Xho*I sites of YIp5-Sc3319 (4a). The deletion endpoints of these derivatives were determined by the DNA sequencing (13) of M13mp9 (11) hybrids containing the relevant *Eco*RI-*Bam*HI fragments. The DNA corresponding to *his3-Δ85* (deleted to −83) was generated by partially cleaving pUC8-Sc2676 DNA (9) with *Dde*I, blunting the ends with DNA polymerase I, ligating an octanucleotide *Sac*I linker (CGAGCTCG), and cloning the 1.4-kbp *Sac*I-*Bam*HI fragment into M13mp19. As expected, the artificially introduced *Sac*I site is immediately adjacent to the *Eco*RI site in the vector polylinker. The *Eco*RI-*Xho*I fragment containing the *his3* structural gene was cloned into YIp5-Sc3319 as above.

Internal deletions were generated by ligating the 5.4-kbp *Bam*HI-*Sal*I vector fragment of YIp55-Sc2812 (18), the 3.7-kbp *Eco*RI-*Sal*I fragments of appropriate 5′ deletion mutations described above, and the 340- to 370-base-pair *Eco*RI-*Bam*HI fragments containing the region from −447 to a position close to the proximal TGACTC sequence. To obtain this last set of DNA fragments, YRp14-Sc2857 (which contains an *Eco*RI linker at position −35) (14) was cleaved with *Eco*RI and treated for various times with *Bal* 31 nuclease. The resulting collection of fragments was treated with DNA polymerase I to blunt the ends, ligated to an *Eco*RI linker, cleaved with *Eco*RI and *Bam*HI, ligated between the *Eco*RI and *Bam*HI sites of M13mp9, and subjected to DNA sequence analysis.

The 1-base-pair deletion of position −99 (*his3-142*) was obtained by extending a 15-nucleotide primer containing

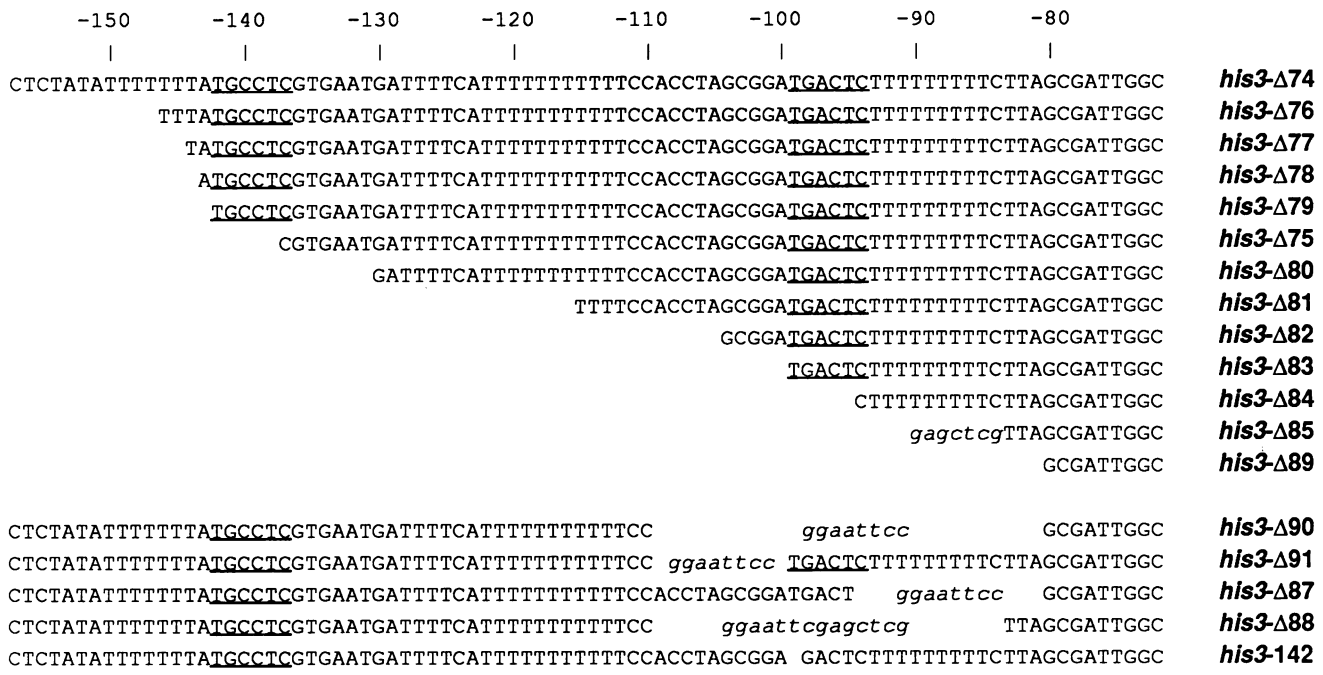

FIG. 2. DNA sequences of *his3* deletion mutations. The *his3* nucleotide sequence of the coding strand between −157 and −72 is shown for each derivative. For the sequential deletions (top), the upstream most nucleotide is joined to the octanucleotide *Eco*RI linker, GGAATTCC. For the internal deletions (bottom), the deleted regions are indicated by spaces, and the linker replacements are shown in small, italic letters.

point mutations at position −99 on M13mp7-Sc2676′ (18). After DNA sequence analysis, we failed to recover any point mutations at this position, but we did obtain the 1-base-pair deletion. The 1.8-kbp *Bam*HI fragment from this derivative was then recloned in the correct orientation into YIp55-Sc2812.

Introduction of deletion DNAs into yeast cells. YIp5 or YIp55 hybrid DNAs were cleaved with *Xba*I and introduced into yeast cells as described previously (18). The transforming DNAs represent the only *his3* alleles in the resulting strains due to the *his3-Δ200* mutation in the hosts. Transformants containing a single copy of the introduced DNA integrated at the *his3* locus were identified by genomic hybridization (data not shown). For the internal deletions and *his3-142*, gene replacement of the *his3-Δ200* allele was performed as described previously (18).

RNA analysis. Three different strains containing a given *his3* deletion allele were grown in YPD broth to an A_{600} of 1.5. Transformants of KY117 were used to measure the constitutive basal level of *his3* expression, transformants of KY466 (*gcd1-1*) were used to measure induced mRNA levels, and transformants of KY498 (*gcn4-2*) were used to analyze transcription in the absence of the GCN4 product. The procedures for isolation of RNA, synthesis and purification of 5′-end-labeled probes, DNA-RNA hybridization, nuclease S1 treatment, and product analysis by gel electrophoresis have been described previously (17, 18). In brief, mRNA from appropriate strains was hybridized to completion with an excess of single-stranded *his3* and *ded1* DNA probes, each labeled at their 5′ ends with ^{32}P. The levels of the *his3* and *ded1* mRNA species were quantitated by scanning the autoradiograms with a DU-6 spectrophotometer (Beckman Instruments, Inc., Fullerton, Calif.). As *ded1* mRNA levels are unaffected by amino acid starvation (17, 20), the *his3* levels shown in Tables 1 and 2 have been normalized to the *ded1* levels.

RESULTS

Sequential deletion analysis. In previous work (17), we generated a series of sequential 5′ deletion mutations which contain the entire *his3* mRNA coding region, the TATA promoter element, and a variable number of base pairs upstream from the TATA region. By joining these *his3* segments to position −447, the resulting molecules were equivalent to the wild-type *pet56-his3-ded1* chromosomal region except that sequences between −447 and a particular point near the *his3* promoter region were deleted (Fig. 1 and 2). The sequences upstream of −447 are located within the *pet56* structural gene and were shown not to influence *his3* transcription. Yeast strains containing single copies of the deletion mutant DNAs at the *his3* locus were grown in broth and analyzed for *his3* mRNA levels. In this way, the constitutive basal levels of *his3* mRNA produced in these derivatives were compared with that of the wild-type *his3* gene (17). The results of these experiments are summarized in Table 1.

In this paper, we analyze these and several new *his3* derivatives for mRNA levels during inducing conditions. The new derivatives, which had deletion endpoints between −137 and −156, were obtained to investigate the role of the second copy of the *his3* regulatory sequence. By standard methods, the deletion DNAs were introduced in a single copy at the *his3* locus of strain KY466 (*ura3-52 his3-Δ200 gcd1-1*). The *gcd1-1* allele in this strain abolishes the translational control of *gcn4* RNA and results in a constitutively high level of GCN4 protein (5, 21). This situation causes induced levels of all coregulated genes even during growth in broth, and it is equivalent to conditions of amino acid starvation (22).

his3 mRNA levels were determined by a quantitative S1 nuclease procedure by using the level of *ded1* RNA as an internal control (17, 18). The observation that *his3* induction

TABLE 1. Sequential deletion mutations

Allele	DNA fragment	Endpoint[a]	*his3* mRNA levels[b] (+1, +12 transcripts)		
			Wild type[c]	*gcn4-2*	*gcdl-1*
HIS3$^+$	Sc2812		1.0, 1.1	0.8, 0.4	1.0, 5.0
his3-Δ72	Sc3403	−192	0.9, 1.0	0.6, 0.3[d]	1.0, 4.8
his3-Δ74	Sc3387	−157	1.0, 1.0	0.6, 0.3	0.8, 4.5
his3-Δ76	Sc3399	−146	NT[e]	NT	0.9, 4.6
his3-Δ77	Sc3400	−144	NT	NT	1.0, 5.2
his3-Δ78	Sc3401	−143	NT	NT	1.1, 5.2
his3-Δ79	Sc3402	−142	NT	NT	1.0, 5.0
his3-Δ75	Sc3365	−137	1.0, 1.2	0.7, 0.3[d]	1.0, 2.4
his3-Δ80	Sc3366	−130	1.0, 1.0	0.7, 0.4	0.9, 2.2
his3-Δ81	Sc3367	−115	0.2, 0.2	NT	0.4, 0.9
his3-Δ82	Sc3368	−104	0.2, 0.2	0.2, 0.2[d]	0.2, 1.0
his3-Δ83	Sc3369	−99	0.2, 0.2	0.2, 0.2	0.3, 1.1
his3-Δ84	Sc3386	−94	<0.1	<0.1	<0.1
his3-Δ85	Sc3370	−83	<0.1	<0.1	<0.1
his3-Δ89	Sc3371	−80	<0.1	<0.1[d]	<0.1[d]

[a] Indicates the nucleotide position at which the *Eco*RI linker was joined.
[b] Normalized to the amount of *ded1* RNA in the same hybridization reaction. The error is approximately ±25%.
[c] The level of the wild-type gene in the wild-type strain was defined to be 1.0; this corresponds roughly to one mRNA molecule per cell (20).
[d] Hybridization data not shown.
[e] NT, Not tested.

is associated with a fivefold increase of the +12 (and minor +22) transcript but no increase of the +1 transcript (19) provides a sensitive method for distinguishing between constitutive and inducible modes of expression. As the relative levels of these transcripts are determined directly, even a twofold increase of the +12 over the +1 transcript (corresponding to a total induction of only 50%) is easily observed. The results of this experiment are shown in Fig. 3 and are summarized below.

Under inducing conditions, deletion mutations that remove sequences as far downstream as −142 (such as *his3-Δ79*) behave indistinguishably from the wild-type *his3* gene. Specifically, the levels of the +1 and +22 transcripts are approximately equal, and the level of the +12 RNA is five times higher. In contrast, derivatives with endpoints at −137 or −130 (*his3-Δ75* and *-Δ80*) do not induce *his3* transcription to the wild-type level. Nevertheless, these deletions confer partial inducibility because the +12 mRNA is two to three times more abundant than the +1 transcript. Deletions *his3-Δ81*, *-Δ82*, and *-Δ83* are also partially inducible, and the absolute levels are about two- to threefold lower. Finally, more extensive deletions (*his3-Δ84* and *-Δ85*) do not have detectable levels of *his3* mRNA. These results indicate that (i) sequences necessary for maximal induction of *his3* transcription are located entirely within the 142 nucleotides adjacent to the mRNA start, (ii) sequences necessary for partial induction are located within the adjacent 104 nucleotides, and (iii) sequences between −142 and −137 and probably −99 and −94 are important for induction.

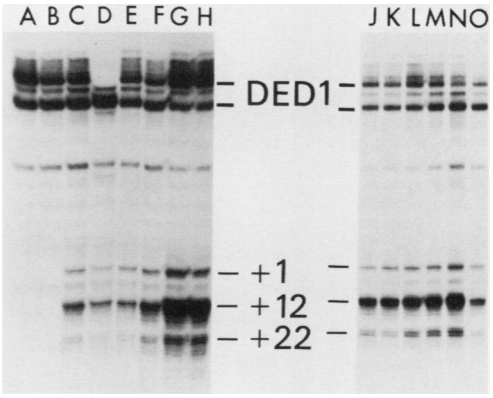

FIG. 3. RNA analysis of sequential deletions. Total RNA (approximately 50 μg) from *gcd1-1* strains (KY466 transformants) containing the following *his3* deletion mutations was hybridized to an excess of single-stranded *his3* and *ded1* probes and treated with S1 nuclease. Lanes: A, *Δ85*; B, *Δ84*; C, *Δ83*; D, *Δ82*; E, *Δ81*; F, *Δ80*; G, *Δ72*; H, wild type; J, *Δ74*; K, *Δ76*; L, *Δ77*; M, *Δ78*; N, *Δ79*; O, *Δ75*. The positions of bands corresponding to the *his3* +1, +12, and +22 transcripts and the *ded1* transcripts are indicated. The band between the *his3* and *ded1* RNAs is due to a trace contamination of double-stranded *his3* probe.

Some of the sequential deletion mutations were also introduced into yeast strain KY498, which contains a *gcn4* mutation and hence is unable to induce transcription in response to amino acid starvation (7, 12). The phenotypic pattern of the various *his3* deletions in the *gcn4-2* strain resembled that observed for constitutive expression (Fig. 4

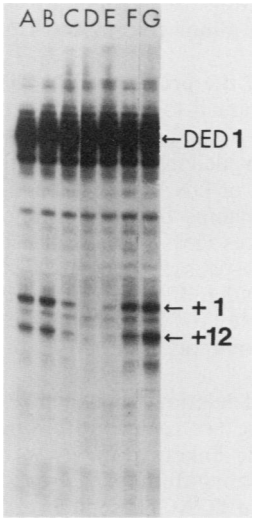

FIG. 4. RNA analysis of sequential deletions in a *gcn4-2* background. Total RNA (approximately 50 μg) from *gcn4-2* strains (KY498 transformants) containing the following *his3* deletion mutations was hybridized to an excess of single-stranded *his3* and *ded1* probes and treated with S1 nuclease. Lanes: A, *Δ74*; B, *Δ80*; C, *Δ83*; D, *Δ84*; E, *Δ85*; F, wild type; G, RNA from strain KY114 (wild-type gene in wild-type background). The positions of bands corresponding to the *his3* +1 and +12 transcripts and the *ded1* transcripts are indicated. The band between the *his3* and *ded1* RNAs is due to a trace contamination of double-stranded *his3* probe.

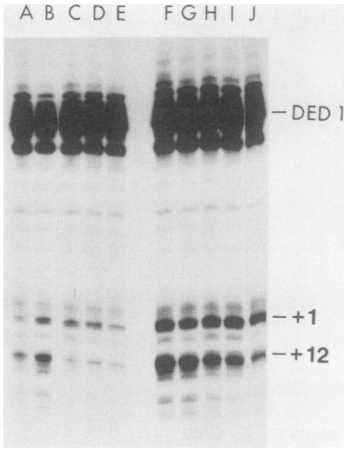

FIG. 5. RNA analysis of internal deletions. RNA (50 μg from wild-type strains and 20 μg from *gcd1-1* strains) containing the following internal deletion mutations was hybridized to an excess of single-stranded *his3* and *ded1* probes and treated with S1 nuclease. Lanes: A and F, Δ*90*; B and G, Δ*91*; C and H, Δ*87*; D and I, Δ*88*; E and J, *his3-142*. Strains represented in lanes A to E contain a *gcd1-1* background (KY466 transformants), and strains represented in lanes F to J contain a wild-type background (KY117 transformants). The positions of bands corresponding to the *his3* +1 and +12 transcripts and the *ded1* transcripts are indicated.

and Table 1). In particular, deletions down to −130 are indistinguishable from the wild-type gene, further deletions down to −99 have approximately 20 to 40% of the normal *his3* mRNA level, and more extensive deletions have less than 10% of the wild-type level. Interestingly, the wild-type gene and some of the deletion mutations seemed to have slightly lower RNA levels in the *gcn4-2* strain as compared with the wild-type strain, and the level of the +12 transcript appeared to be preferentially reduced by a small factor with respect to the +1 transcript. This may suggest that GCN4 protein may play a minor role in constitutive *his3* expression.

Small deletions of the proximal TGACTC element abolish inducibility. The original localization of the *his3* regulatory element between −83 and −103 was based on the phenotypes resulting from deletion mutations that removed relatively large regions of DNA and hence significantly altered the spacing relationships between promoter elements and mRNA initiation sites (15). The smallest deletion that abolished *his3* induction was 53 base pairs. Although spacing alterations generally had little effect on *his3* expression (14, 15), it was desirable to analyze the phenotypes of small deletions that remove or disrupt the proximal TGACTC element.

Four new internal deletions, in which 10 to 30 base pairs of the native *his3* gene are replaced by 8-base-pair *Eco*RI or *Sac*I oligonucleotide linkers (Fig. 2), were obtained as described in Materials and Methods. Three of these derivatives destroy the proximal TGACTC sequence, and one derivative (*his3-Δ91*) replaces 10 base pairs just upstream of the proximal TGACTC with an *Eco*RI linker. As described above, the mutant DNAs were introduced into wild-type and *gcd1-1* strains such that they precisely replaced the normal *his3* chromosomal locus. The resulting strains were grown in broth and analyzed for *his3* mRNA levels. The results are shown in Fig. 5 and summarized in Table 2.

In a wild-type genetic background, i.e., normal conditions, all five internal deletion mutations behaved indistinguishably from the native *his3* gene. However, in a *gcd1-1* background, these mutations were unable to induce *his3* expression above the basal level. The *his3-Δ91* allele, which replaces sequences just adjacent to the proximal TGACTC may have slightly elevated levels of *his3* mRNA under these conditions. Thus, small deletions of or near the proximal TGACTC do not affect the basal level of *his3* transcription, but they abolish inducibility under conditions of amino acid starvation.

Single-base-pair deletion at position −99 prevents *his3* induction. In our attempts to create point mutations of the proximal TGACTC sequence, we initially employed the primer extension method described by Zoller and Smith (23). For various reasons, some of which we do not understand, these attempts were unsuccessful. However, in the course of these experiments, we did obtain a 1-base-pair deletion of the thymine residue at −99 (*his3-142*), which changed the TGACTC sequence to AGACTC (Fig. 2). After replacement of the *his3* locus, *his3* mRNA levels conferred by the *his3-142* allele were examined as described above. This mutation had little, if any, effect on the basal level of expression, and it failed to induce *his3* transcription in response to starvation conditions (Fig. 5).

DISCUSSION

Two copies of the element are necessary and sufficient for maximal induction. The observation that *his3-Δ79*, which deletes sequences upstream of −142, behaves indistinguishably from the wild-type gene indicates that the two proximal copies of the TGACTC sequence are sufficient for maximal *his3* induction. Conversely, this result indicates that the four distal copies are unnecessary for induction. Thus, these distal copies appear to be nonfunctional derivatives of the *his3* regulatory element, although it is possible that they could be functional in other contexts.

The two regulatory sequences present in *his3-Δ79* include a perfect (TGACTC) copy between −99 and −94 and an imperfect (TGCCTC) copy between −142 and −137. The imperfect copy is important for maximal induction because *his3-Δ75*, which is identical to *his3-Δ79* except for the deletion of five additional residues, does not fully induce *his3* transcription. The perfect copy is absolutely required for induction because small deletions, including a 1-base-pair deletion, abolish inducibility. This result also indicates that the imperfect copy, together with the four far upstream copies, is inactive in the absence of the perfect copy. In contrast, the perfect copy is sufficient for partial induction

TABLE 2. Internal deletion mutations

Allele	DNA fragment	Endpoint[a]	*his3* mRNA levels[b] (+1, +12 transcripts)	
			Wild type[c]	*gcd1-1*
HIS3$^+$	Sc2812		1.0, 1.0	1.0, 5.0
his3-Δ90	Sc3260	−109/−80	1.0, 1.2	0.8, 1.1
his3-Δ91	Sc3261	−109/−99	0.8, 1.0	1.1, 1.5
his3-Δ87	Sc3269	−95/−80	1.0, 0.8	1.0, 0.7
his3-Δ88	Sc3384	−109/−83	1.1, 0.9	1.0, 0.9
his3-142	Sc3401	−100/−98	0.9, 0.6	0.8, 0.7

[a] Nucleotide positions between which the natural *his3* sequence was deleted. The deleted region in *his3-Δ90*, *-Δ91*, and *-Δ87* was replaced by an *Eco*RI linker, and the deleted region in *his3-Δ88* was replaced by an *Eco*RI and *Sac*I linker.

[b] As in Table 1, footnote *b*.

[c] As in Table 1, footnote *c*.

because deletions down to −99 show an increase in the level of +12 transcription during starvation conditions.

Our sequential deletion mapping of the *his3* regulatory sequences yields qualitatively similar results to the analyses of the *his4* promoter. Maximal levels of *his4* induction are achieved with a promoter containing a proximal TGACTC element centered at position −140 and two closely spaced distal elements centered at −185 and −200 (1). Deletions that remove both distal copies but retain the proximal copy are inducible, but not to the maximal level. In accord with analogous *his3* deletions, some of these derivatives also reduce the basal level of *his4* expression.

One apparent difference between *his3* and *his4* induction seems to be the requirement for the proximal copy. Unlike the *his3* situation where very small deletions of the proximal copy abolish inducibility, relatively large deletions of the *his4* proximal copy that retain both distal copies are still inducible (8). However, the absolute level of transcription in such derivatives is 50- to 100-fold below the maximal induction level (8). Our interpretation of these results is that the *his4* distal copies are extremely defective regulatory sequences, whose effects can only be seen in large deletions that remove the elements for constitutive transcription. We predict that these sequences will be bound very weakly by GCN4 protein, which would explain both the low levels of expression and the inducibility. Although both *his4* distal copies are perfect TGACTC sequences unlike the imperfect *his3* distal sequence, the adjacent residues in the distal *his4* copies are poorly matched to the consensus for regulatory sequences found in 15 amino acid biosynthesis genes (4a). In contrast, both the *his3* and *his4* proximal copies are well matched to the consensus, and by themselves are capable of induction levels that are only two- to three- fold below the maximal level.

Thus, it appears that the *his3* and *his4* promoters are organized in a very similar manner. In both cases, at least two regulatory sequences are necessary for maximal induction levels, but the proximal copies are clearly more important. One difference between these promoters, however, concerns the sequences for constitutive transcription. Although the sequences for constitutive *his4* transcription have yet to be defined, this promoter does not contain a poly(dA-dT) region similar to that which is critical for the basal level of *his3* expression (17).

Functional distinctions between the proximal and distal regulatory sequences. Repeated sequences in eucaryotic promoter regions are often essential for the maximal level of transcription, and in some cases they have been shown to be multiple binding sites for a given protein (2, 3). Presumably, the effects on transcription involve an increase in the local concentration of recognition sites that facilitates protein binding or cooperative interactions between bound proteins.

However, several observations are suggestive of functional distinctions between the proximal and distal *his3* regulatory sequences. Unlike the proximal copy which is sufficient for partial induction, the distal copy is inactive by itself. Moreover, in previous work we demonstrated strong binding of GCN4 protein to the proximal copy and no detectable binding to the distal copy (or any of the other four copies) (9). The DNase I footprinting experiments indicated that the affinity of GCN4 protein for the distal copy is at least 50-fold lower than that of the proximal copy, and they showed no evidence for cooperation.

The imperfection of the distal copy, TGCCTC instead of TGACTC, is likely to account for its relative inactivity. In the presence of all five upstream copies, a 1-base-pair substitution that changes the proximal copy to TGCCTC abolishes induction in vivo and GCN4 protein binding in vitro (4a). Nevertheless, despite the apparently defective nature of the distal copy, the results in this paper clearly indicate that it is important for achieving the maximal level of *his3* transcription during inducing conditions.

This apparent anomaly can be explained in two ways. One possibility is that in contradiction to our DNA-binding experiments (9), GCN4 protein does bind the distal copy in vivo, possibly by using cooperative interactions that we have been unable to detect in vitro. An alternative explanation, which we favor, is that an additional DNA-binding protein recognizes similar DNA sequences and binds to the distal copy. This latter suggestion is reminiscent of the model for *cyc1* activation in which related DNA sequences confer distinct regulatory patterns and respond to different *trans*-acting factors (4).

While direct evidence for this hypothesis is lacking, several observations are suggestive. First, *his3-Δ83* confers a low basal level of expression that is independent of the poly(dA-dT) element and the wild-type *gcn4* allele. Since *his3-Δ84*, a deletion of 5 additional base pairs that removes the proximal TGACTC, fails to confer this residual basal level of transcription, this suggests that the proximal TGACTC sequence acts as a weak upstream regulatory element in a *gcn4*-independent manner. Second, although *his3-Δ83* and *his3-Δ82* have indistinguishable phenotypes, *his3-Δ83* DNA is weakly bound by GCN4 protein in vitro (4a), unlike *his3-Δ82* DNA which is bound as well as the native gene (9). This may mean that sequences immediately upstream of −99 are important for GCN4 action, but may be less important for binding by a separate factor. Third, in derivatives lacking sequences upstream of −102 and hence the poly(dA-dT) element and the imperfect distal copy, *his3* expression is increased by a mutation that changes the proximal sequence to TGACTT (19). However, the identical mutation abolishes *his3* induction in the presence of all five distal copies (4a). One interpretation of these observations is that specific changes within a common recognition sequence could differentially affect binding by GCN4 and the hypothetical protein. Thus, these functional distinctions between the native proximal and distal copies and between mutational forms of the proximal copy suggest that the TGACTC sequence may interact with two distinct proteins.

ACKNOWLEDGMENTS

We thank Ian Hope and Sarah Bray for fruitful discussions.

D.E.H. was supported by postdoctoral grants from the Damon Runyon and Massachusetts Medical Foundation. This work was funded by Public Health Service grant GM30186 from the National Institutes of Health and by the Chicago Community Trust (Searle Scholars Program).

LITERATURE CITED

1. **Donahue, T. F., R. S. Daves, G. Lucchini, and G. R. Fink.** 1983. A short nucleotide sequence required for the regulation of *his4* by the general control system of yeast. Cell **32**:89–98.
2. **Gidoni, D., J. T. Kadonaga, H. Barrera-Saldana, K. Takahashi, P. Chambon, and R. Tjian.** 1985. Bidirectional SV40 transcription mediated by tandem Sp1 binding interactions. Science **230**:511–517.
3. **Giniger, E., S. M. Varnum, and M. Ptashne.** 1985. Specific DNA binding of GAL4, a positive regulatory protein of yeast. Cell **40**:767–774.
4. **Guarente, L., B. Lalonde, P. Gifford, and E. Alani.** 1984. Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *S. cerevisiae*. Cell

36:503–511.

4a. **Hill, D. E., I. A. Hope, J. P. Macke, and K. Struhl.** 1986. Saturation mutagenesis of the yeast *his3* regulatory site: requirements for transcriptional induction and for binding by the GCN4 activator protein. Science **234**:451–457.

5. **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.

6. **Hinnebusch, A. G., and G. R. Fink.** 1983. Repeated DNA sequences upstream from *his1* also occur at several other co-regulated genes in *Saccharomyces cerevisiae*. J. Biol. Chem. **258**:5238–5247.

7. **Hinnebusch, A. G., and G. R. Fink.** 1983. Positive regulation in the control of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **80**:5374–5378.

8. **Hinnebusch, A., G. Lucchini, and G. R. Fink.** 1985. A synthetic *HIS4* regulatory element confers general amino acid control on the cytochrome C gene (*CYC1*) of yeast. Proc. Natl. Acad. Sci. USA **82**:498–502.

9. **Hope, I. A., and K. Struhl.** 1985. GCN4 protein, synthesized in vitro, binds to *his3* regulatory sequences: implications for the general control of amino acid biosynthetic genes in yeast. Cell **43**:177–188.

10. **Jones, E. W., and G. R. Fink.** 1982. Regulation of amino acid biosynthesis and nucleotide biosynthesis in yeast, p. 181–299. *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.

11. **Messing, J., and J. Vieira.** 1982. A new pair of M13 vectors for selecting either DNA strand of double digest restriction fragments. Gene **19**:269–276.

12. **Penn, M. D., B. Galgoci, and H. Greer.** 1983. Identification of *AAS* genes and their regulatory role in general control of amino acid biosynthesis in yeast. Proc. Natl. Acad. Sci. USA **80**:2704–2708.

13. **Sanger, F., A. R. Coulson, B. G. Barell, A. J. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.

14. **Struhl, K.** 1982. The yeast *his3* promoter contains at least two distinct elements. Proc. Natl. Acad. Sci. USA **79**:7385–7389.

15. **Struhl, K.** 1982. Regulatory sites for *his3* expression in yeast. Nature (London) **300**:284–287.

16. **Struhl, K.** 1983. Direct selection for gene replacement events in yeast. Gene **26**:231–241.

17. **Struhl, K.** 1985. Naturally occurring poly dA:dT sequences are upstream promoter elements for constitutive transcription in yeast. Proc. Natl. Acad. Sci. USA **82**:8419–8423.

18. **Struhl, K.** 1985. Nucleotide sequence and transcriptional mapping of the yeast *pet56-his3-ded1* gene region. Nucleic Acids Res. **13**:8587–8601.

19. **Struhl, K., W. Chen, D. E. Hill, I. A. Hope, and M. A. Oettinger.** 1985. Constitutive and coordinately regulated transcription of yeast genes: promoter elements, positive and negative regulatory sites, and DNA binding proteins. Cold Spring Harbor Symp. Quant. Biol. **50**:489–503.

20. **Struhl, K., and R. W. Davis.** 1981. Transcription of the *his3* gene region in *Saccharomyces cerevisiae*. J. Mol. Biol. **152**:535–552.

21. **Thireos, G., M. D. Penn, and H. Greer.** 1984. 5′ untranslated sequences are required for the translational control of a yeast regulatory gene. Proc. Natl. Acad. Sci. USA **81**:5096–5100.

22. **Wolfner, M., D. Yep., F. Messenguy, and G. R. Fink.** 1975. Integration of amino acid biosynthesis into the cell cycle of *Saccharomyces cerevisiae*. J. Mol. Biol. **96**:273–290.

23. **Zoller, M. J., and M. Smith.** 1982. Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. Nucleic Acids Res. **10**:6487–6500.