MOLECULAR AND CELLULAR BIOLOGY, Feb. 1988, p. 655–663
0270-7306/88/020655-09$02.00/0

Vol. 8, No. 2

The *HAP3* Regulatory Locus of *Saccharomyces cerevisiae* Encodes Divergent Overlapping Transcripts

STEVEN HAHN, JENNIFER PINKHAM,† RANDY WEI, REBECCA MILLER, AND LEONARD GUARENTE*

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received 27 August 1987/Accepted 5 November 1987

Activation of the *CYC1* upstream activation site, UAS2, and transcription of several other genes encoding respiratory functions requires the product of the regulatory gene *HAP2*. We report here the isolation and characterization of a second UAS2 regulatory gene, *HAP3*. Like mutations in *HAP2*, a mutation in *HAP3* abolishes the activity of UAS2 and prevents growth on nonfermentable carbon sources. The *HAP3* gene was cloned and, surprisingly, was found to encode two divergently transcribed, overlapping transcripts: a 570-base RNA and a 3-kilobase (kb) RNA. Chromosomal disruption experiments defined the critical region for *HAP3* function to a 1.3-kb segment in which the two transcripts overlap. Analysis of the *HAP3* DNA sequence showed that the 570-base transcript could encode a protein of 144 amino acids. Synthesis of the 144-amino-acid protein under regulatory control in vivo demonstrated that this protein is essential for activity of UAS2 as well as for growth on nonfermentable carbon sources. The largest open reading frame in the critical region of the 3-kb transcript is only 86 amino acids. Using site-directed mutagenesis, we demonstrated that the 86-amino-acid open reading frame was not involved in UAS2 activity. The possible role of this 3-kb antisense RNA in HAP3 expression or function is discussed.

An understanding of eucaryotic gene regulation requires an analysis of regulatory proteins at a biochemical level. In animal cells, such biochemical studies involve the identification and purification of protein factors from extracts which bind to particular regulatory sequences and, in some cases, activate transcription in vitro (for example, see references 4, 16, 29, 36). In *Saccharomyces cerevisiae* studies of regulatory proteins have proceeded along different lines. Initially, genes that encode regulatory proteins were identified by mutations that altered the regulation of particular structural genes. Such loci include *GAL4* (15, 18, 27), *GCN4* (12), and *HAP1* (10), which are positive regulators, and *MATα2* (20, 39), a negative regulator. The products of all these genes have been shown to bind to their cognate regulatory sites in vitro (3, 8, 13–15, 17, 30).

In some cases, *GCN4*, for example, analysis of cloned sequences has provided insight into how regulation occurs. Amino acid biosynthetic genes under *GCN4* control derepress when cells are starved for amino acids (general control) (12). This regulation is due to an increase in the rate of synthesis of *GCN4* itself that occurs at the transcriptional level (26). A long leader in the *GCN4* mRNA contains regulatory sequences that result in a low level of translation of *GCN4* under repressed conditions and a high level of translation under amino acid starvation conditions.

In this paper, we report new findings in our study of the regulation of genes encoding cytochromes and related proteins in *S. cerevisiae*. The *CYC1* gene encoding iso-1-cytochrome *c* contains tandem independent upstream activation sites, UAS1 and UAS2 (10, 11). UAS2 is highly regulated by carbon catabolite repression and derepresses about 50-fold when cells are shifted from a glucose medium to one containing a nonfermentable carbon source. In cells lacking normal levels of intracellular heme, the basel level of UAS2 is greatly reduced (10). The activity of UAS2 is abolished by mutations in a locus, *HAP2* (10, 31). *HAP2* mutant strains are pleiotropically deficient in cytochromes. Experiments in our laboratory have shown that the transcription of at least two genes, *COX4* (nuclear cytochrome oxidase subunit 4) and *HEM1* (δ-aminolevulinate synthase), is abolished in such a mutant (J. Schneider, T. Keng, and L. Guarente, unpublished data). Thus, *HAP2* is a global regulator of respiratory functions in yeast cells (31). The product of *HAP2* is a 265-amino-acid nuclear protein, the carboxyl third of which is highly basic (32). Further studies indicate that a HAP2-lexA fusion protein can stimulate transcription when bound to the lexA operator in vivo, thus raising the possibility that HAP2 is a direct activator of transcription.

Data presented herein suggest that *HAP2* itself is insufficient to activate UAS2 or genes encoding other respiratory functions. We describe a mutation in a second locus, *HAP3*, with properties identical to those of mutations in *HAP2*. We also describe the isolation, sequencing, and transcriptional mapping of the *HAP3* region. Because of the complexity in the transcriptional map of the *HAP3* locus, a detailed mutational analysis was done to discern *HAP3* functional sequences. This analysis showed that a 144-amino-acid protein encoded at *HAP3* is functional and left open possible roles of an RNA that is transcribed from the opposite strand and overlaps the mRNA encoding the 144-amino-acid protein.

MATERIALS AND METHODS

Strains and plasmids. *S. cerevisiae* strains used in this work are all derivatives of BWG1-7a (*Mata leu2-2,2-112 his4-519 ade1-100 ura3-52*) (11). The UAS2-*lacZ* fusion Δ229 up-1 (10) was integrated at the chromosomal *LEU2* locus as described previously (26) in both BWG1-7a and a *hap3-1* strain to create strains SHY25 and SHY27, respectively. *Escherichia coli* MC1061 (F^- *araD139* Δ(*ara-leu-7697*) Δ*lac-74* $GalU^-$ $GalK^-$ $HsdR^-$ $HsdM^+$ Str^r) and JM101 were used for all DNA manipulations in vitro.

The yeast genomic DNA library was provided by K. Pfeifer and consisted of a *Sau*3A1 digest of total DNA from strain BWG1-7a cloned at the *Bam*HI site of the single-copy

* Corresponding author.
† Present address: Section of Molecular Neurobiology, Yale School of Medicine, New Haven, CT 06510.

ARS CEN yeast vector YCP50. All *HAP3* fragments were subcloned into either YCP50 or the high-copy plasmid YEP24 (5).

Assays and medium. β-Galactosidase assays were performed as described previously (8). Cells were grown in minimal medium consisting of yeast nitrogen base without amino acids and without ammonium sulfate supplemented with 2% carbon source, 40 μg of required amino acids per ml, and 0.004% adenine before growth.

General methods. Standard yeast genetic procedures were followed (37). All standard DNA manipulations were as described previously (21). Total yeast RNA was isolated as described in reference 41. Poly(A)$^+$ RNA was selected with Hybond-mAP polyuridylic acid paper (Amersham Corp., Arlington Heights, Ill.) following the method described by the manufacturer.

Isolation and cloning of the *hap3-1* mutant. Strain BWG1-7a bearing the UAS2-*lacZ* fusion plasmid pLGΔ229 up-1 (10) was mutagenized with ethyl methanesulfonate to 50% survival and screened for β-galactosidase levels on XG indicator plates as described previously (10). Of the 60,000 colonies screened, 9 appeared pale blue on indicator plates compared with controls and showed reduced levels of β-galactosidase in liquid assays. These strains were cured of plasmid Δ229 up-1 and retransformed with this plasmid as well as the UAS1-*lacZ* fusion plasmid pLGΔAlu-Xho (10) and pLGSD5, a UAS_{GAL}-*lacZ* fusion plasmid (10). Only one mutant strain, JP60, was specifically reduced for only UAS2 transcription. This strain bore a nuclear petite mutation, *hap3-1*, which segregated 2:2 in crosses with wild type. The *HAP3* gene was cloned by transforming this mutant strain with a yeast genomic library and selecting for complementation of the petite phenotype. Of the six clones analyzed, two unique overlapping clones were isolated as described below.

Chromosomal disruption of *HAP3*. The 2.8-kilobase (kb) *Hin*dIII fragment of *HAP3* was cloned into the yeast integration vector YIP5 (40) at the *Hin*dIII site to create plasmid pSH102. For disruption at the *HAP3 Xho*I site, the *Xho*I-*Sal*I *LEU2* fragment from plasmid YCP13 (37) was cloned into the *Xho*I site of pSH102. This construct was cut with *Mlu*I and *Pst*I and transformed to strain BWG1-7a, selecting for Leu$^+$ recombinants which had inserted the *LEU2* fragment at the chromosomal *HAP3* locus. For disruption of HAP3 with the *HIS4* gene, the *HIS4 Sac*I-*Bst*EII fragment of pPB54 (7; J. Fink, personal communication) was inserted into plasmid pSH102 at the *Cla*I site, the *Pvu*II site, or between the *Xho*I-*Sac*I sites. These plasmids were cut with *Mlu*I and *Pst*I and used to transform strain SHY25 (*his4 LEU2*::UAS2-*lacZ*), selecting for His recombinants which had inserted the *HIS4* fragment at the *HAP3* locus. The expected structure of all chromosomal disruptions was confirmed by Southern analysis of chromosomal DNA (data not shown).

Northern (RNA) blots. Total cellular RNA was fractionated by size and transferred to nitrocellulose as described previously (21). RNA probes were generated with SP6 RNA polymerase (22). Hybridization to RNA probes was done at 63°C overnight in 50% formamide–5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) (21)–3.5× Denhardt (21)–50 mM sodium phosphate (pH 6.5)–0.1% sodium dodecyl sulfate–250 μg of calf thymus DNA per ml–250 μg of tRNA per ml. Filters were washed in 0.2× SSC–0.1% sodium dodecyl sulfate at 65°C.

S1 mapping. S1 mapping was done as described previously (2). S1 probes were made by copying M13 single-stranded clones with DNA polymerase in the presence of [α-^{32}P] dATP. Total yeast RNA was hybridized to single-stranded M13 probes in 80% formamide–48 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] (pH 6.4)–0.48 M NaCl–1.2 mM EDTA at 45°C for 3 h. A 350-μl portion of 5% glycerol–2 mM $ZnCl_2$–0.1 M NaCl–60 mM sodium acetate–6,000 U of S1 nuclease (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) per ml was added and incubated for 25 min at 37°C followed by analysis of products on 6% acrylamide sequencing gels.

DNA sequencing. Restriction fragments of the *HAP3* gene were cloned into the M13 vectors mp18 and mp19 and sequenced by the dideoxynucleotide chain termination method of Sanger et al. (34).

Mapping the *HAP3* lesion by recombination. Strain SHY27 (*hap3-1 LEU2*::UAS2-*lacZ*) containing either plasmid pSH84 (*HAP3 Xho*I-*Eco*RV fragment [Fig. 1] in plasmid YCP50) or pSH87 (*HAP3 Eco*RI-*Xho*I fragment [Fig. 1] in plasmid YCP50) was plated on rich lactate plates at 10^7 cells per plate. Neither of those two plasmids complements *hap3-1* for growth on lactate medium. After 2 days at 30°C, stable lactate-positive cells appeared from the strain containing pSH84 at a frequency of 10^{-6}. No lactate-positive revertants were seen with the strain containing pSH87. When cured of pSH84, the lactate-positive recombinants grew like wild type on lactate and had wild-type levels of UAS2 expression.

Cloning the *hap3-1* mutation. The *hap3-1* mutation was cloned by transforming the *hap3-1* strain with the plasmid pSH94 (see Fig. 7) cut with *Mlu*I and *Pvu*II and treated with alkaline phosphatase. Ura$^+$ *hap3* colonies were picked, and plasmid DNA was isolated and analyzed. One of eight clones examined contained the chromosomal *hap3-1* DNA inserted correctly in the *Mlu*I-*Pvu*II gap created in plasmid pSH94.

Fusion of the 144-amino-acid ORF to *lacZ*. Plasmid pSH97 (containing the *HAP3 Hin*dIII fragment in plasmid YCP50) was cut with *Hin*dIII and *Ban*I at positions 1 and 2020, respectively (see Fig. 6), and made flush with DNA polymerase. Synthetic 10-mer *Bam*HI linkers (New England BioLabs, Inc., Beverly, Mass.) were ligated to this DNA followed by digestion with *Bam*HI. The resulting 2-kb *Bam*HI *HAP3* fragment was isolated. The *CYC1-lacZ* fusion plasmid pLGΔ312 (10) containing a *Bgl*II linker at *Sma*I (B. Lalonde, unpublished data) was digested with *Bgl*II and *Bam*HI. This plasmid backbone was ligated with the *HAP3 Bam*HI fragment to create a fusion of the 144-amino-acid open reading frame (ORF) to *lacZ* (pSH127). The expected structure was confirmed by DNA sequencing of the fusion junction. Synthesis of the 144-amino-acid ORF was placed under control of either the *CYC1* UASs or the *GAL* UAS by cutting pSH127 with *Xho*I and filling the sticky ends with DNA polymerase. This DNA was cut with *Sac*I in a partial digest, and the 3-kb *Xho*I-*Sac*I fragment containing *HAP3-lacZ* was isolated. This fragment was inserted into plasmid backbones derived from either pLG669-ATG or SD5-ATG (9) which were cut with *Bam*HI and *Sac*I. The resulting plasmids contain the *HAP3-lacZ* fusion protein under control of the *CYC1* TATA and mRNA initiation region and either the *CYC1* UAS or the *GAL* UAS (pSH151 and pSH152, respectively).

Site-directed mutagenesis of the 86-amino-acid *HAP2* ORF. The 800-base-pair (bp) *HAP3 Sac*I-*Bam*HI fragment was cloned into pEMBL19 (6) and used to generate single-stranded DNA by infection with bacteriophage F1 as described previously (6). This DNA was annealed to 0.5 μg of a kinased oligonucleotide of sequence AACGGGGAAGA TATTCTTATATCATTGCACGC in 80 mM Tris hydrochlo-

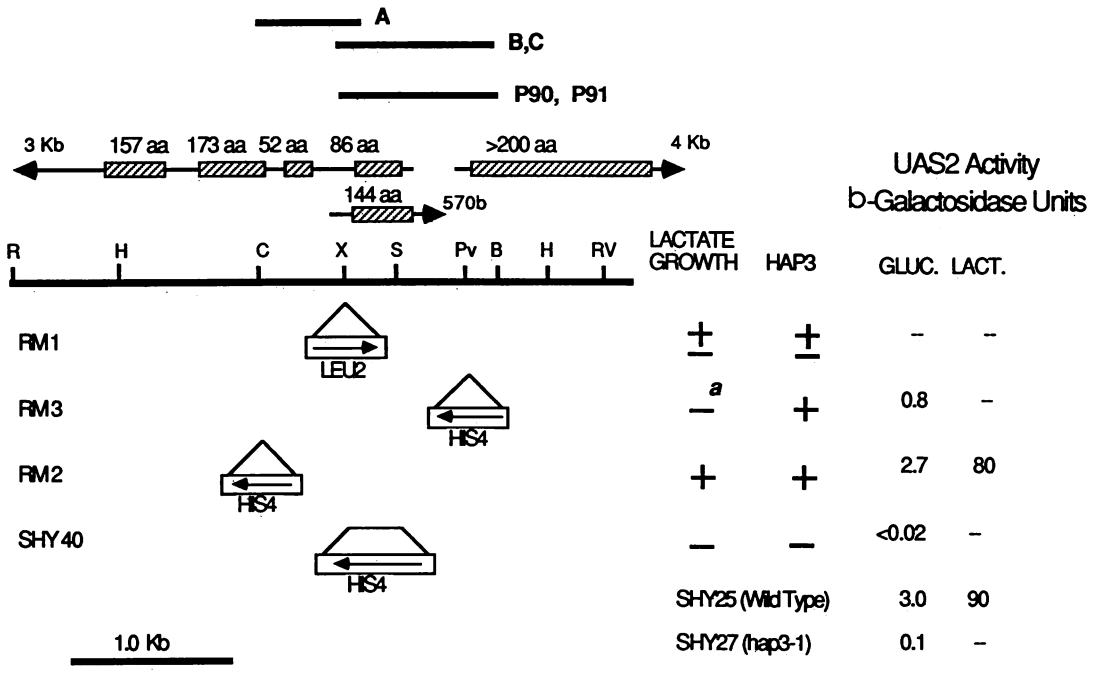

FIG. 1. Chromosomal gene disruptions at the *HAP3* locus. The positions of *LEU2* and *HIS4* gene fragments inserted into the chromosomal *HAP3* locus in strain BWG1-7a (RM1) or SHY25 (RM2, RM3, SHY40) are indicated. Lactate growth phenotype was measured on rich lactate plates, and *HAP3* complementation was determined by mating with a *hap3-1* strain and measuring growth of the diploid on rich lactate. β-Galactosidase activity was measured in minimal medium containing the indicated carbon source (GLUC., glucose; LACT., lactate). Arrows represent transcripts encoded at the *HAP3* locus, and hatched boxes indicate ORFs encoded within the transcripts (see text). A, B, and C represent single-stranded DNA probes used in S1 mapping experiments. P90 and P91 represent single-stranded RNA probes used in Northern analysis. R, *Eco*RI; C, *Cla*I; X, *Xho*I; S, *Sac*I; Pv, *Pvu*II; RV, *Eco*RV; B, *Bam*HI; H, *Hin*dIII. [a], RM3 had a severe growth defect on all media, growing very slowly on glucose medium and not at all on lactate medium. aa, Amino acids; b, bases.

ride (pH 7.6)–16.7 mM $MgCl_2$–21 mM NaCl–0.2 mM EDTA. After hybridization was complete, 8.5 μl of 600 μM each dATP, dGTP, dCTP, and dTTP, 0.6 mM ATP, 12 mM dithiothreitol, 0.6 U of DNA polymerase Klenow (Boehringer Mannheim) per μl, and 24 U of T4 DNA ligase (New England BioLabs) per μl was added and incubated at 15°C for 1 h followed by incubation at room temperature overnight (23).

This DNA was used to transform *E. coli* JM101 to Amp^r. Single-stranded DNA was made from 48 transformants and screened for the desired double point mutation by differential hybridization with the 32-mer oligonucleotide (23). Two positive candidates were analyzed by DNA sequencing and found to contain the desired mutations. The DNA was used to reconstruct an intact *HAP3* gene by excising the mutant *HAP3 Sac*I-*Pvu*II fragment and using it to replace the wild-type sequence in both pSH94 and pSH95 (see Fig. 6).

RESULTS

Isolation of the *hap3-1* mutation. UAS2 regulatory mutants were isolated as detailed in Materials and Methods. Of the 60,000 mutagenized colonies screened, 1 mutant strain (JP60) was found which reduced expression of a *CYC1-lacZ* fusion driven by UAS2 about 40-fold (reduced from 24 U of β-galactosidase in wild type to 0.65 U) but had no effect on expression of the UAS elements UAS1 (38 U in wild type and 50 U in JP60) and UAS_{GAL} (4,000 U in wild type and 3,000 U in JP60). Strain JP60 also failed to grow on nonfermentable carbon sources and showed lower levels of transcription from *COX4* (cytochrome oxidase subunit 4) as well as *HEM1*, which encodes a heme biosynthetic enzyme (J. Schneider, T. Keng, and L. Guarente, unpublished data). All the above phenotypes are identical to those caused by mutation of the *HAP2* gene (10, 31). The mutation in this strain, *hap3-1*, segregated as a single nuclear locus in a genetic cross. In 20 tetrads examined, the petite phenotype segregated 2:2. In the four tetrads tested, both the petite and the UAS2-deficient phenotypes segregated together. Further, a diploid strain of *hap2-1/hap3-1* showed complementation of both UAS2 expression and growth on nonfermentable carbon sources, showing that the mutations lie in distinct genes. Sporulation of the diploid and tetrad dissection showed that *HAP2* and *HAP3* are unlinked (data not shown).

Cloning of the *HAP3* gene. As detailed in Materials and Methods, the *HAP3* gene was cloned by transforming strain JP4-5c (*hap3-1 LEU2*::UAS2-*lacZ*) (11) with a single-copy yeast genomic library and selecting for complementation of the *hap3-1* petite phenotype. Two unique clones were isolated which fully complemented the lactate-negative phenotype of strain JP4-5c and restored activation of UAS2 as measured in a UAS2-*lacZ* fusion integrated at the *LEU2* locus.

To demonstrate that these clones contained the *HAP3* gene and not a suppressor of the *hap3-1* mutation, we used the cloned DNA to create a series of gene disruptions at the chromosomal *HAP3* locus (Fig. 1). As detailed in Materials and Methods, DNA fragments containing either the *LEU2* or *HIS4* gene were integrated into the chromosome by the method of Rothstein (33a). Cells with an insertion of *LEU2* at the *HAP3 Xho*I site (RM1) showed poor growth on lactate as well as sevenfold-lower levels of UAS2 expression (Table 1). Replacement of the *HAP3 Xho*I-*Sac*I region by *HIS4* (SHY40) prevented growth on lactate and reduced UAS2

TABLE 1. Activity of UAS2 in *hap3*::*LEU2* disruption strain RM1[a]

Strain	β-Galactosidase (U)	
	Glucose	Lactate
BWG1-7a (wild type)	59	520
JP60 (*hap3-1*)	<2	
RM1	8.5	111

[a] The indicated strains were transformed with the UAS2-*lacZ* fusion plasmid pLGΔ265 up-1 (11) and assayed for β-galactosidase in minimal medium containing the indicated carbon source.

activity about 100-fold as measured in a strain containing a chromosomal UAS2-*lacZ* fusion. Both strains RM1 and SHY40 were completely complemented when transformed with a single-copy plasmid containing the 2.8-kb *Hin*dIII fragment. Neither the lactate growth defect nor UAS2 expression in these strains was complemented by mating with a strain containing the *hap3-1* mutation, thus demonstrating that we cloned the authentic *HAP3* gene.

Insertion of *HIS4* at the *HAP3* *Cla*I site (RM2) had no effect on either lactate growth or UAS2 activity. Insertion of *HIS4* at *Pvu*II (RM3) resulted in an unusual phenotype we attributed to effects on a neighboring gene: a severe growth defect on rich glucose medium as well as an inability to grow on nonfermentable carbon sources. This strain was complemented for both wild-type growth and UAS2 activity by mating with a *hap3-1* strain. This showed that the phenotype of RM3 is due to the disruption of a gene distinct from *HAP3*. Unlike the disruptions in *HAP3*, this strain was not complemented by the single-copy subclone of the 2.8-kb *Hin*dIII fragment. Full complementation of this strain required additional DNA greater than 3 kb to the right of the *Pvu*II site. From these and other data presented below, we conclude that strain RM3 is defective not in *HAP3* but in another neighboring gene. This strain showed a small reduction in levels of UAS2 expression (Fig. 1). However, this effect appears nonspecific to UAS2 since a *HIS4-lacZ* fusion is also reduced about fourfold in this strain (data not shown). We believe that this phenotype is likely due to the very poor growth of this strain and tentatively conclude that this gene neighboring *HAP3* plays no direct role in UAS2 expression. The above gene disruption analysis delineates the region essential for *HAP3* function as the 1.3 kb between the *Cla*I and *Pvu*II sites.

***HAP3* encodes two transcripts.** To characterize transcripts encoded by the *HAP3* locus, we constructed probes by cloning the *Xho*I-*Bam*HI *HAP3* fragment adjacent to the SP6 promoter in both orientations (Fig. 1). Radioactively labeled RNA was synthesized from each construct and was hybridized to cellular RNA which had been size fractionated on formaldehyde agarose gels and transferred to nitrocellulose. Probe P90, detecting RNAs transcribed from right to left (Fig. 1), hybridized to a 3-kb transcript (Fig. 2A). This RNA is synthesized constitutively in glucose and lactate media, in the *hap2-1* and *hap3-1* mutant strains, and in cells deficient in heme biosynthesis (data not shown). Probe P91, which hybridizes to transcripts synthesized in the opposite orientation from P90, detected two transcripts, one of 570 nucleotides (migrating with an apparent size of 800 bases) and another of about 4 kb (Fig. 2B). Again, levels of these RNAs were not affected by carbon source or mutations in *HAP2* or *HAP3*.

Northern analysis of poly(A)-enriched RNA showed that the 570-base RNA is polyadenylated, as 50 to 90% was reproducibly retained by selection with poly(U) paper (Fig. 2B). Surprisingly, 80 to 90% of the 3-kb and 4-kb transcripts were reproducibly lost upon a single round of poly(A) selection (Fig. 2A and B). We are not certain whether this was due to a difficulty in selecting large polyadenylated mRNA by our method (see Materials and Methods).

The fact that all three of these transcripts were encoded near the *HAP3* locus was confirmed by analyzing the pattern of transcripts produced in strains containing *HIS4* or *LEU2* chromosomal disruptions of *HAP3* (Fig. 3). Strain RM1

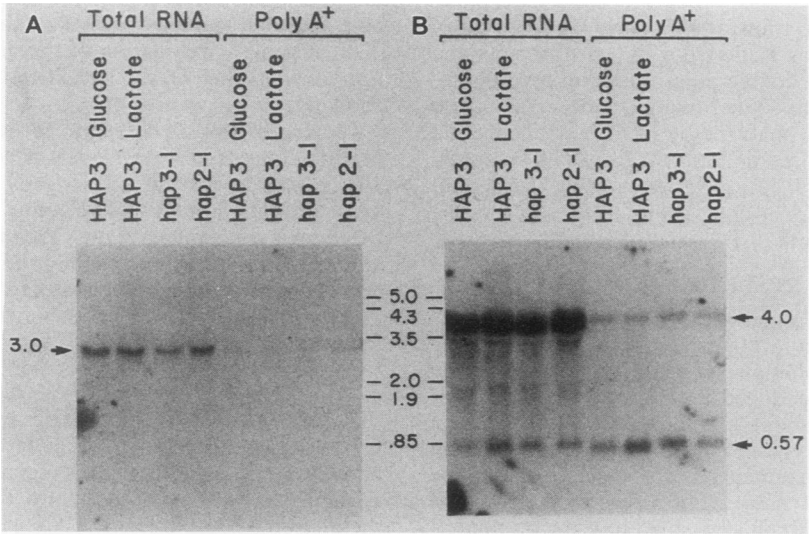

FIG. 2. Northern analysis of RNAs encoded at the *HAP3* locus. Northern blots of either 30 μg of total cellular RNA or poly(A)$^+$ RNA extracted from 60 μg of total RNA probed with single-stranded *Xho*I-*Bam*HI fragment (Fig. 1). (A) Probe P90 detects RNAs transcribed from right to left (Fig. 1). (B) Probe P91 detects RNAs transcribed from left to right. RNA was extracted from indicated strains grown in minimal glucose (*hap3-1*, *hap2-1*) or minimal lactate medium as indicated. Numbers indicate size in kilobases.

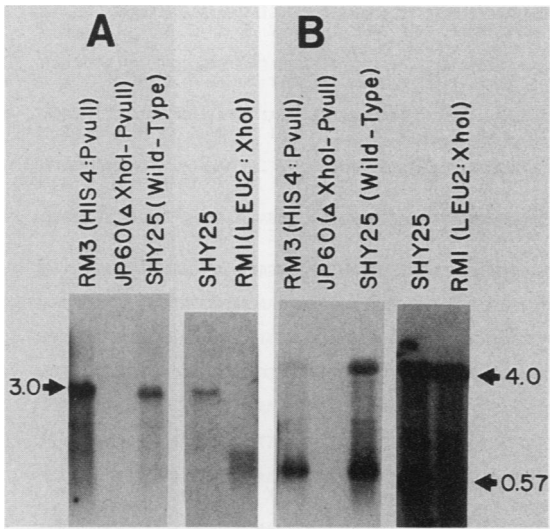

FIG. 3. Effect of *HAP3* gene disruptions on RNAs encoded by *HAP3*. Shown are Northern blots of 30 μg of total cellular RNA from indicated strains grown in minimal glucose medium. (A) Probe P90; (B) Probe P91. Numbers indicate size in kilobases.

containing a *LEU2* insertion at *Xho*I of *HAP3* produced a truncated 3-kb transcript (Fig. 3A) and no detectable 570-base transcript (Fig. 3B). This disruption did not affect the synthesis of the 4.0-kb transcript, suggesting that this RNA corresponded to the neighboring gene and not *HAP3*. Consistent with this proposal, insertion of *HIS4* at *Pvu*II (RM3) eliminated the 4-kb transcript, but did not affect transcription of either the 3-kb or 570-base transcript (Fig. 3). Strain JP60, which contains a replacement of the chromosomal *Xho*I-*Pvu*II sequences by *HIS4* (J. Pinkham, unpublished data), showed no detectable RNAs with either probe P90 or P91 (Fig. 3). Finally, insertion of *HIS4* at *Cla*I (RM2) had no effect on the 570-base transcript but did increase the size of the 3-kb transcript owing to transcription across the *HIS4* insert (data not shown).

The above data show that the *HAP3* locus encodes two overlapping transcripts of 570 bases and 3 kb that are read off of opposite strands (Fig. 1). The 4-kb RNA detected in this region corresponds to the gene that lies to the right of *HAP3* and is transcribed away from *HAP3* (Fig. 1).

Mapping the 5′ and 3′ ends of the *HAP3* transcripts. We wished to determine the extent of overlap of the two *HAP3* RNAs. Thus, probes were prepared to S1 map the 5′ end of the 3-kb RNA and both the 5′ and 3′ ends of the 570-base RNA (Fig. 1; Materials and Methods). The 5′ end of the 570-base RNA was mapped to two major start sites about 14 and 31 bp to the left of the *Xho*I site (Fig. 4A). The 3′ end of this transcript was placed about 530 bp to the right of *Xho*I (Fig. 4B). The 5′ end of the 3-kb RNA was placed about 480 bp to the right of *Xho*I (Fig. 4C). These data indicate that the 3-kb RNA starts 40 bp from the 3′ end of the 570-base RNA and is read from the opposite strand (Fig. 1). Additional information gleaned from this mapping is that the *LEU2*::*Xho*I insertion does not abolish left-to-right transcription across the *HAP3* locus (asterisk, Fig. 4A). This transcription could explain the leaky phenotype of this mutation. These experiments also place the 5′ end of the 4-kb RNA about 680 bases to the right of the *Xho*I site (Fig. 4B). All the above RNA endpoints are indicated in the *HAP3* sequence figure (see below).

DNA sequence of *HAP3*. The DNA sequence of the 3.1-kb *Hin*dIII-*Eco*RV fragment was determined by using overlapping M13 clones and the dideoxy sequencing method (Fig. 5). Analysis of the DNA sequence showed that the 570-base RNA could encode a protein of 144 amino acids with the first initiator AUG codon in the message being that of the ORF. This protein contains about 10% basic and 15% acidic residues. It does not contain obvious homology to either the helix-turn-helix motif common to several DNA-binding proteins (28, 35) or the zinc finger domains common to several other known DNA-binding proteins (24, 33).

Analysis of potential protein-coding regions in the 3-kb transcript indicated the largest ORF in the critical *Xho*I-*Pvu*II region to be of 86 amino acids. This ORF overlaps the 144-amino-acid ORF in the 570-base RNA. The initiator AUG codon of this ORF is the second one found from the 5′ end of the RNA, however.

The DNA sequence encoding the 4-kb transcript has the

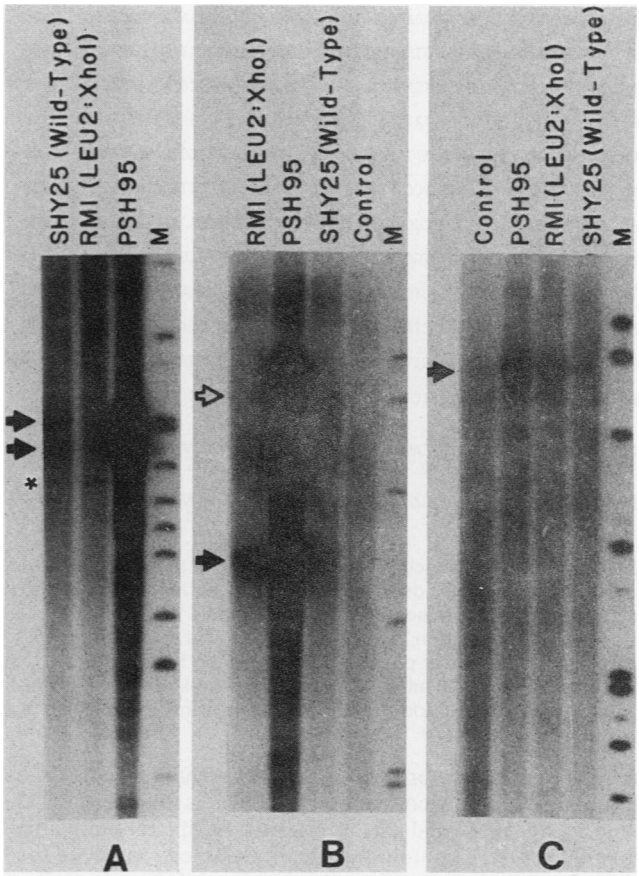

FIG. 4. Mapping the 5′ and 3′ ends of transcripts encoded at the *HAP3* locus. Endonuclease S1 analysis of total cellular RNA. (A) Probe A is single-stranded *Cla*I-*Sph*I fragment (positions 1020 to 1800) (Fig. 1 and 6) in an orientation that detects RNAs transcribed from left to right (Fig. 1). (B) Probe B is single-stranded *Xho*I-*Bam*HI fragment (Fig. 1) that detects RNAs transcribed from left to right (Fig. 1). (C) Probe C is single-stranded *Xho*I-*Bam*HI fragment that detects RNAs transcribed from right to left (Fig. 1). All reactions contain 20 μg of total cellular RNA from the indicated strains grown in minimal glucose medium, except pSH95 in panel C which used 5 μg. pSH95 is a high-copy *HAP3* plasmid grown in a wild-type strain (see Fig. 6). Control contains no RNA; lane M is a size standard of *Msp*I-cut pBR322. Filled arrows point to 5′ RNA ends, and open arrow points to 3′ end of 570-base RNA. Asterisk marks truncated 570-base RNA seen in strain RM1 (*HAP3*:*LEU2*).

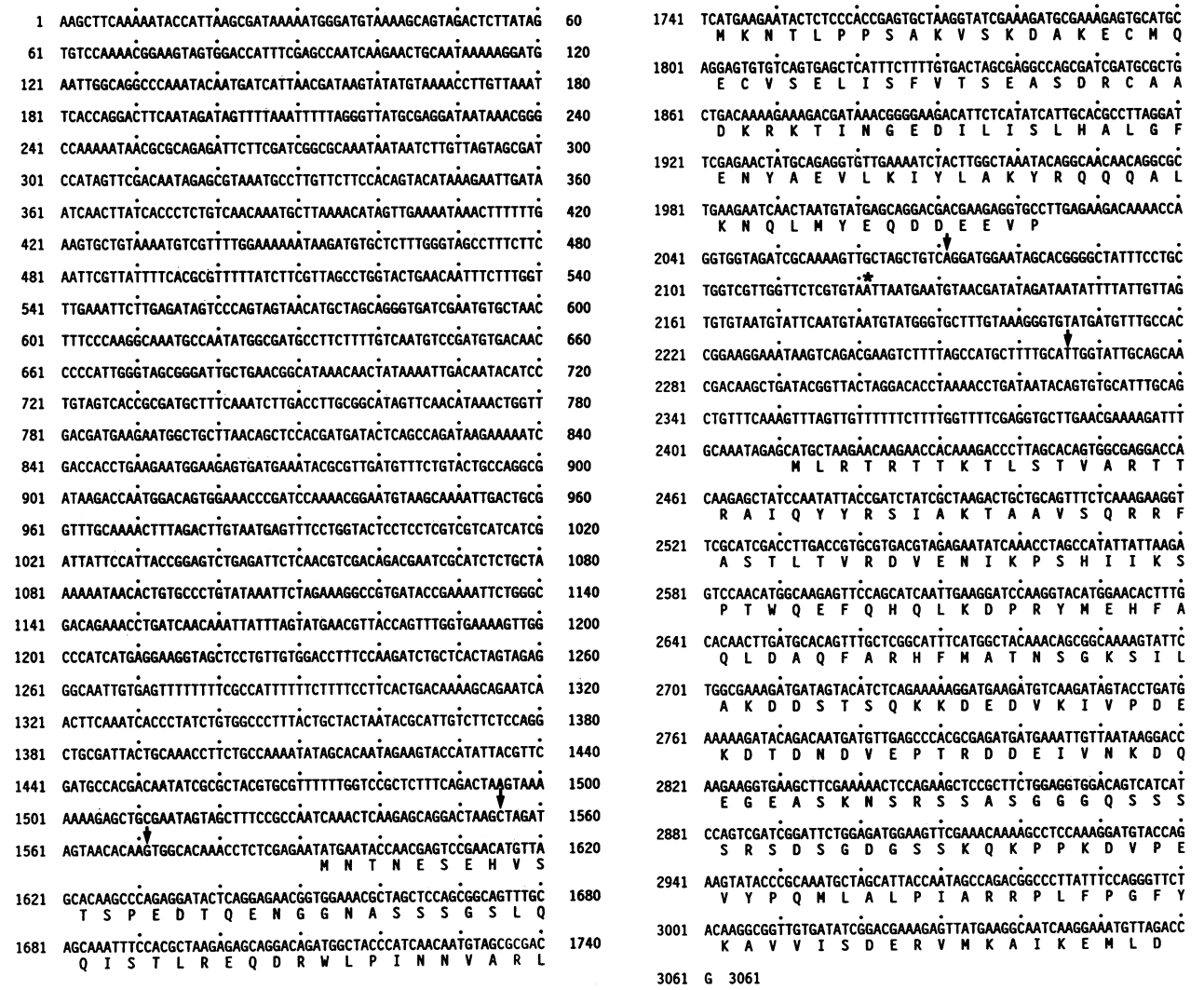

FIG. 5. DNA sequence of *HAP3*. Shown is the DNA sequence of the 3,061-bp *Hin*dIII-*Eco*RV fragment (Fig. 1). The sequence is numbered from the leftmost *Hin*dIII site (Fig. 1). The sequence was determined on both DNA strands except for the region between 2870 and 3061 which was determined on only one strand. Arrowheads at positions 1555 and 1571 mark the 5′ end of the 570-base transcript ± 5 bp, and an asterisk at position 2121 marks the 3′ end of this transcript ± 5 bp. Arrowhead at position 2071 indicates the 5′ end of the 3-kb transcript ± 5 bp, and arrowhead at position 2267 marks the 5′ end of the 4-kb transcript ± 5 bp. Also shown are the 144-amino-acid ORF of the 570-base transcript and the >200-amino-acid ORF of the 4-kb transcript. The 86-amino-acid ORF of the 3-kb transcript lies between positions 1900 and 1657. The *hap3-1* mutation is a single G-to-A transition at position 1884.

potential to encode a protein of at least 200 amino acids. The initiator methionine codon of this ORF is the first AUG codon from the 5′ end of the 4-kb transcript. None of the above three ORFs contain significant homology to any of the protein sequences in either the National Biomedical Research Foundation protein data base or the Doolittle protein data base.

Sequence of the *hap3-1* mutation. We wished to determine whether *HAP3* function was encoded in the 144-amino-acid protein or the 3-kb RNA. To begin, the position of the *hap3-1* mutation was mapped by using recombination to rescue the *hap3-1* mutant phenotype (see Materials and Methods). A single-copy subclone which contains *HAP3* DNA only to the right of *Xho*I could not complement the lactate growth defect of *hap3-1* but could recombine with the *hap3-1* mutant to generate stable *HAP3*$^+$ recombinants. This maps the *hap3-1* lesion to the right of *Xho*I. As detailed in Materials and Methods, the chromosomal sequences from *Pvu*II to *Mlu*I were recombined onto the single-copy plasmid pSH94 containing the 3.4-kb *Hin*dIII-*Eco*RI fragment (Fig. 6). The *hap3-1* lesion was further mapped by subcloning the mutant *Cla*I-*Eco*RI fragment into plasmid pSH94 deleted from *Cla*I-*Eco*RI. The resulting plasmid (pRW3) did not complement a *hap3-1* strain (Fig. 6). The *hap3-1* lesion was also inserted into a high-copy yeast vector by subcloning the *hap3-1 Xho*I-*Pvu*II fragment into pSH95 (Fig. 7) to generate pRW4. In high copy, *hap3-1* allowed weak growth on lactate and activated UAS2 to about 20% of its normal level in lactate medium. This demonstrates that *hap3-1* still retains some *HAP3* activity. Sequencing of the 900-bp *Xho*I-*Bam*HI fragment showed that the *hap3-1* lesion is a single G→A transition at position 1884 (Fig. 7). This mutation results in a glycine-to-arginine change at position 98 of the 144-amino-acid ORF and a proline-to-leucine change at position 6 in the 86-amino-acid ORF.

The 144-amino-acid ORF is required for activity of UAS2.

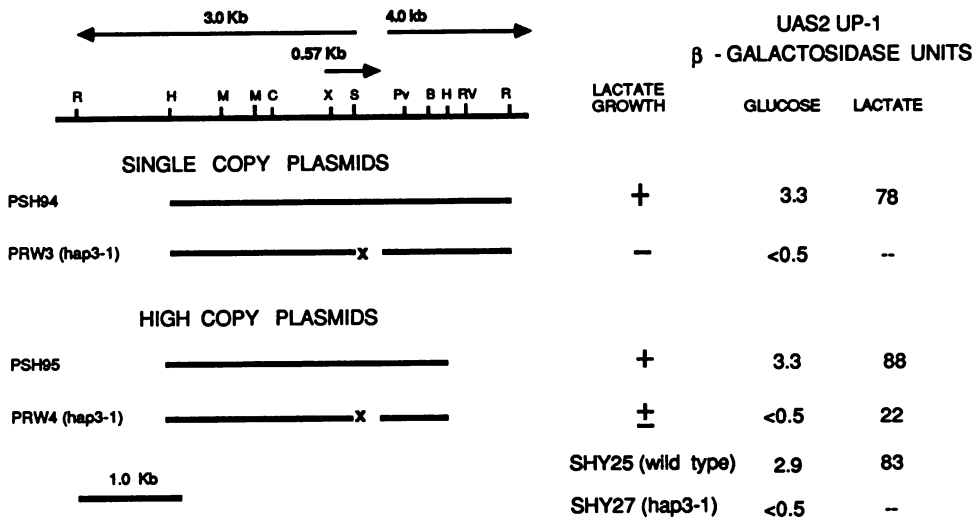

FIG. 6. The *hap3-1* allele partially activates UAS2 when present in high copy. Strain SHY27 (*hap3-1* UAS2-*lacZ*::*LEU2*) was transformed with the indicated plasmids. Lactate growth was measured on rich lactate plates, and β-galactosidase was measured in minimal medium with the indicated carbon source. R, *Eco*RI; H, *Hin*dIII; C, *Cla*I; X, *Xho*I; S, *Sac*I; PV, *Pvu*II; B, *Bam*HI; RV, *Eco*RV; M, *Mlu*I.

The sequence of the *hap3-1* mutation still did not distinguish which RNA encoded *HAP3* function. To approach this question, we placed the 144-amino-acid ORF under regulatory control. To monitor the synthesis of the 144-amino-acid protein, we fused the 144th codon of the ORF to the *E. coli lacZ* gene, resulting in a bifunctional fusion protein (see Materials and Methods). The fusion was then inserted into plasmids which placed protein synthesis under control of either the *CYC1* UAS1 and UAS2 elements or the UAS_{GAL} element. The 144-amino-acid–*lacZ* fusion produced 0.5 U of β-galactosidase activity when synthesized under control of the wild-type *HAP3* promoter (Fig. 8). This plasmid partially complemented both the *hap3-1* mutant and a *HAP3* deletion strain, SHY41 (Δ*Xho*I-*Sac*I::*HIS4*), for growth on lactate as well as restoring about 25% of wild-type UAS2 activity. When placed under the control of *CYC1*, the fusion protein produced 270 U of β-galactosidase activity in lactate medium and fully complemented both the *hap3-1* mutant and the Δ*HAP3* strain SHY41 for growth on lactate. When placed under control of the UAS_{GAL} (which is inactive in the absence of galactose), the fusion produced less than 0.03 U of β-galactosidase activity and did not complement *hap3-1* or Δ*HAP3* for either lactate growth or UAS2 activity in the absence of galactose. Galactose induced β-galactosidase expression to 230 U. The above results demonstrated that synthesis of the 144-amino-acid ORF is essential for activation of UAS2 and complementation of the lactate deficiency of the *hap3* deletion. All three of the above *lacZ* fusion plasmids, however, also adventitiously synthesized RNA starting in *lacZ* and transcribed in an antisense orientation across the 144-amino-acid ORF (data not shown). This antisense transcript could potentially direct the synthesis of the 86-amino-acid ORF normally encoded on the 3-kb *HAP3* RNA.

To determine whether the synthesis of the 86-amino-acid ORF was also essential for UAS2 activity, we constructed a double point mutation which eliminated the first two methionine codons at positions 1 and 3 in the 86-amino-acid ORF (see Materials and Methods) (Fig. 9). The next methionine codon in the 86-amino-acid ORF is at position 27. The site-directed mutations were constructed so as not to change the 144-amino-acid ORF synthesized in the opposite orientation. When inserted into either pSH94 (single copy) or pSH95 (high copy), the mutant *HAP3* gene completely complemented both *hap3-1* and Δ*HAP3* strains for UAS2 activity and for lactate growth. From this, we conclude that the 86-amino-acid ORF plays no role in the activity of UAS2.

DISCUSSION

In this report, we describe the identification of a locus, *HAP3*, required for the activation of UAS2 of the *CYC1* gene and for the global activation of genes encoding cytochromes and related products. The phenotype of *HAP3* mutants is identical to that of *HAP2* mutants: an inability to grow on nonfermentable carbon sources.

There are two general mechanisms that could explain the requirement for both *HAP2* and *HAP3*. Either one gene could regulate the expression or activity of the other gene product or, alternatively, both gene products might function together in a complex. Two observations are consistent with the latter possibility. First, experiments presented herein and described previously showed that the expression of *HAP2* and *HAP3* does not depend on functionality of the other locus (31, 32; S. Hahn, unpublished data). Second, defects at either the *HAP2* and *HAP3* locus are not bypassed by overexpression of the other gene product (31; S. Hahn, unpublished data). Biochemical experiments are in progress to test whether *HAP2*, *HAP3*, or a complex containing both will bind to UAS2 in vitro.

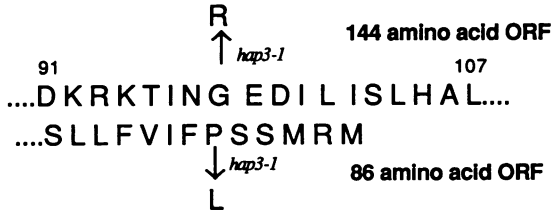

FIG. 7. The *hap3-1* mutation changes the sequence of both the 144-amino-acid and the 86-amino-acid ORFs. The top line depicts the amino acid sequence of the 144-amino-acid ORF from positions 91 to 107, and the bottom line shows the sequence of the 86-amino-acid ORF from positions 1 to 13. The sequence changes caused by the *hap3-1* mutation are indicated by arrows.

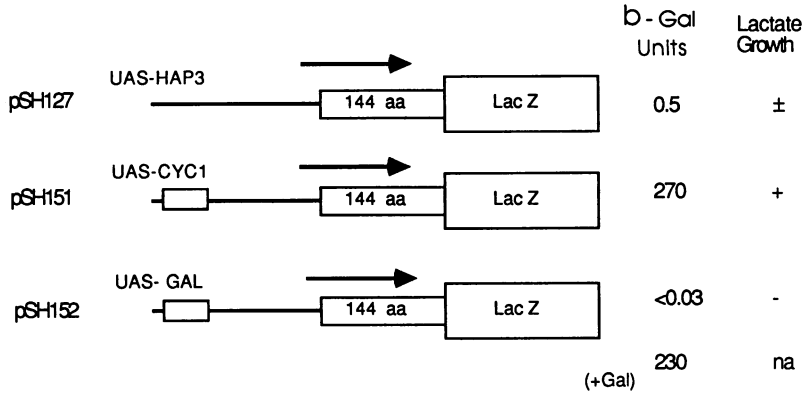

FIG. 8. Synthesis of the 144-amino-acid ORF is essential for *HAP3* activity. Shown are HAP3-lacZ fusion proteins under control of the *HAP3* UAS, the *CYC1* UASs, or the *GAL* UAS. Plasmids were transformed into strain SHY41 (Δ*HAP3*:*HIS4*), and β-galactosidase (b-Gal) was measured in minimal lactate (pSH127, pSH151) or minimal glucose (pSH152) medium. Lactate growth was measured on rich lactate plates. na, Not applicable; +Gal, grown in minimal galactose medium.

Analysis of the cloned *HAP3* locus showed that extensively overlapping transcripts are encoded. Although the methods used here cannot precisely determine the relative amounts of the two transcripts, both Northern and S1 analysis suggested that the two RNAs are synthesized in roughly equivalent amounts. One transcript is 570 nucleotides long and encodes a protein of 144 amino acids. Fusion of this ORF to that encoding β-galactosidase showed that, indeed, it is translated. The AUG that starts this ORF is the first such triplet of the transcript. The other transcript of 3 kb is transcribed in the opposite direction, initiating close to the 3′ end of the 570-nucleotide RNA. This RNA contains no sizable ORF. Chromosomal insertion mutations and the sequencing of the *hap3-1* allele indicated that *HAP3* function is encoded in the overlap region of the two RNAs. By placing the 144-amino-acid ORF under control of the *CYC1* or *GAL1-10* promoters and examining the phenotype of a *HAP3* deletion strain bearing such constructs, we demonstrated that this protein is crucial to *HAP3* function.

What is the role of the 3-kb RNA? The longest ORF encoded by this transcript in the region of overlap with the 570-base RNA is 86 codons. By oligonucleotide-directed mutagenesis, we abolished the start codon for this ORF in a manner that created a silent change in the 144-codon ORF. This mutant complemented a *HAP3* deletion strain, showing that the 86-codon ORF is not involved in *HAP3* function. We are left with at least three possible explanations for the role of the 3-kb RNA. First, it could be a structural RNA that is required for the activity of an activation complex containing HAP2, HAP3, and possibly other proteins. Second, it could regulate expression of the 144-amino-acid protein, perhaps by forming an RNA-RNA hybrid with the 570-base RNA to downregulate the translatability of this message. Such antisense regulation has been observed in several procaryotic systems (1, 19, 25, 38, 42) but has not yet been observed in any naturally occurring eucaryotic system. Third, the 3-kb RNA may be involved in a function unrelated to *HAP3*, or may be fortuitous. We consider this last possibility unlikely because of the intimate juxtaposition of this transcript with that encoding *HAP3*. Experiments are in progress to test the role of the 3-kb RNA in *HAP3* expression or function.

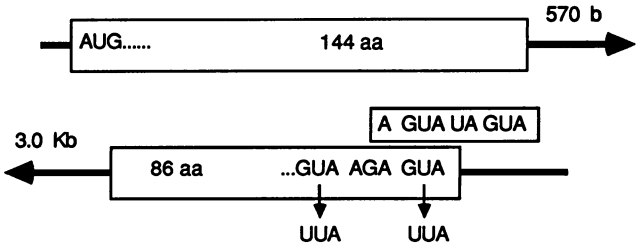

FIG. 9. The 86-amino-acid (aa) ORF is not required for *HAP3* activity. The open boxes indicate the overlap between the 144-amino-acid and 86-amino-acid ORFs. The 3-kb RNA contains a two-codon ORF just upstream from the start of the 86-amino-acid ORF as indicated. Arrows show the double point mutations generated by site-directed mutagenesis. These changes did not change the amino acid sequence of the 144-amino-acid ORF. b, Bases.

ACKNOWLEDGMENTS

We thank K. Pfeifer for help with *HAP3* cloning; A. Brunelle and J. Tobin for advice on site-directed mutagenesis; members of the Guarente laboratory for invaluable discussions throughout the course of the work; and Clea A. Zolotow for help in preparing the manuscript.

This work was supported by a Damon Runyon-Walter Winchell Cancer Fund postdoctoral fellowship to S.H., a postdoctoral fellowship from the National Institutes of Health to J.P., and Public Health Service grant 5 5 ROI GM30454-05 from the National Institutes of Health to L.G.

LITERATURE CITED

1. **Aiba, H., S. Matsuyama, and M. S. Mizunot.** 1987. Function of *micF* as an anti-sense RNA in osmoregulatory expression of the *ampF* gene in *Escherichia coli*. J. Bacteriol. **169**:3007–3012.
2. **Berk, A. J., and P. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
3. **Bram, R. J., and R. D. Kornberg.** 1985. Specific protein binding to far upstream activating sequences in polymerase II promoters. Proc. Natl. Acad. Sci. USA **82**:43–47.
4. **Briggs, M. R., T. T. Kadonaga, S. P. Bell, and R. Tjian.** 1986. Purification and biochemical characterization of the promoter-specific transcription factor, SP1. Science **234**:47–52.
5. **Carlson, M., and D. Botstein.** 1982. Two differently regulated mRNAs with different 5′ ends encode secreted and intracellular forms of yeast invertase. Cell **28**:145–154.
6. **Dente, L., G. Cesareni, and R. Cortese.** 1983. PEMBL: a new family of single stranded plasmids. Nucleic Acids Res. **11**:1645–1655.
7. **Donahue, T. F., P. J. Farabaugh, and G. R. Fink.** 1982. The

nucleotide sequence of the HIS4 region of yeast. Gene 18:47–59.
8. **Giniger, E., S. M. Varnum, and M. Ptashne.** 1985. Specific DNA binding of *GAL4*, a positive regulatory protein of yeast. Cell 40:767–774.
9. **Guarente, L.** 1983. Yeast promoters and *lacZ* fusions designed to study expression of cloned genes in yeast. Methods Enzymol. 101:181–191.
10. **Guarente, L., B. Lalonde, P. Gifford, and E. Alani.** 1984. Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *Saccharomyces cerevisiae*. Cell 36:503–511.
11. **Guarente, L., and T. Mason.** 1983. Heme regulates transcription of the *CYC1* gene of *Saccharomyces cerevisiae* via an upstream activation site. Cell 32:1279–1286.
12. **Hinnebusch, A. G., and G. R. Fink.** 1983. Positive regulation in the general amino acid control of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA 80:5374–5378.
13. **Hope, I., and K. Struhl.** 1985. GCN4 protein, synthesized *in vitro*, binds *HIS4* regulatory sequences: implications for the general control of amino acid biosynthetic genes in yeast. Cell 43:177–188.
14. **Johnson, A. D., and I. Herskowitz.** 1985. A repressor (*MAT*a2 product) and its operator control expression of a set of cell type specific genes in yeast. Cell 42:237–247.
15. **Johnston, S., and J. Hopper.** 1982. Isolation of the yeast regulatory gene *GAL4* and analysis of its dosage effects on the galactose/melibiose regulon. Proc. Natl. Acad. Sci. USA 79: 6971–6975.
16. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter *in vitro*. Cell 42:559–572.
17. **Keegan, L., G. Gill, and M. Ptashne.** 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science 231:699–704.
18. **Laughlon, A., and R. Gesteland.** 1982. Isolation and preliminary characterization of the *GAL4* gene, a positive regulator of transcription in yeast. Proc. Natl. Acad. Sci. USA 79:6827–6831.
19. **Liao, S., T. Wu, C. Chian, M. M. Susskind, and W. McClure.** 1987. Control of gene expression in bacteriophage P22 by a small anti-sense RNA. I. Characterization *in vitro* of the P_{sar} promoter and the *Sar* RNA transcript. Genes Dev. 1:197–203.
20. **MacKay, V., and T. R. Manney.** 1974. Mutations affecting sexual conjugation and related processes in *Saccharomyces cerevisiae*. II. Genetic analysis of non-mating mutants. Genetics 76:273–288.
21. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Melton, D., P. Kreig, T. Rebagliati, T. Maniatis, T. Kinn, and M. Green.** 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing an SP6 promoter. Nucleic Acids Res. 18:7035–7056.
23. **Middleton, S. A., and E. R. Kantrowitz.** 1986. Importance of the loop at residues 230-245 in the allosteric interactions of *Escherichia coli* aspartate carbamoyltransferase. Proc. Natl. Acad. Sci. USA 83:5866–5870.
24. **Miller, J., A. McLachlan, and A. Klug.** 1985. Repetitive zinc binding domains in the protein transcription factor TFIIIA from *Xenopus* oocytes. EMBO J. 4:1609–1654.
25. **Mizuno, T., M. Chen, and M. Inone.** 1984. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA (micRNA). Proc. Natl. Acad. Sci. USA 81:1966–1970.
26. **Mueller, P. P., and A. G. Hinnebusch.** 1986. Multiple upstream AUG codons mediate translational control of GCN4. Cell 45:201–207.
27. **Oshima, Y.** 1982. Regulatory circuits for gene expression: the metabolism of galactose and phosphate, p. 159–180. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), Molecular biology of the yeast *Saccharomyces cerevisiae*: metabolism and gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Pabo, C. O., and M. Lewis.** 1982. Operator binding domain of lambda repressor. Nature (London) 298:443–447.
29. **Parker, C. G., and J. Topol.** 1984. A *Drosophila* RNA polymerase II transcription factor binds to the regulatory site of an hsp70 gene. Cell 37:273–283.
30. **Pfeifer, K., B. Arcangioli, and L. Guarente.** 1987. Yeast *HAP1* activator competes with the factor RC2 for binding to the upstream activation site UAS1 of the *CYC1* gene. Cell 49:9–18.
31. **Pinkham, J., and L. Guarente.** 1985. Cloning and molecular analysis of the *HAP2* locus: a global regulator of respiratory genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 5:3410–3416.
32. **Pinkham, J., J. Olesen, and L. Guarente.** 1987. Sequence and nuclear localization of the yeast HAP2 protein, an activator of transcription. Mol. Cell. Biol. 7:578–585.
33. **Rosenberg, U., C. Schroder, A. Preiss, K. Rienlin, S. Cote, I. Riede, and A. Jackle.** 1986. Structural homology of the product of the Drosophila Kruppel gene with *Xenopus* transcription factor IIIA. Nature (London) 319:336–340.
33a. **Rothstein, R.** 1983. One-step gene disruption in yeast. Methods Enzymol. 101:202–211.
34. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.
35. **Sauer, R., R. Yocum, R. Doolittle, M. Lewis, and C. Pabo.** 1982. Homology among DNA binding proteins suggests use of a conserved super secondary structure. Nature (London) 298: 447–450.
36. **Sawadogo, M., and R. G. Roeder.** 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell 43:165–175.
37. **Sherman, F., G. R. Fink, and J. B. Hicks.** 1983. Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
38. **Simons, R. N., and N. Kleckner.** 1983. Translational control of IS10 transposition. Cell 34:683–691.
39. **Strathern, J., J. Hicks, and I. Herskowitz.** 1981. Control of cell type in yeast by the mating type locus: the $alpha_1$-$alpha_2$ hypothesis. J. Mol. Biol. 147:357–372.
40. **Struhl, K., D. Stinchcomb, S. Scherer, and R. Davis.** 1979. High frequency transformation of yeast: autonomous replication of hybrid DNA molecules. Proc. Natl. Acad. Sci. USA 76:1035–1039.
41. **Teem, J., and M. Rosbash.** 1983. Expression of a β-galactosidase gene containing the ribosomal protein 51 intron is sensitive to the rna2 mutation of yeast. Proc. Natl. Acad. Sci. USA 80:4403–4407.
42. **Tomizawa, J., and T. Itoh.** 1981. Plasmid ColE1 incompatibility determined by interaction of RNA1 with primer transcript. Proc. Natl. Acad. Sci. USA 78:6098–6100.