MOLECULAR AND CELLULAR BIOLOGY, Aug. 1989, p. 3231–3243
0270-7306/89/083231-13$02.00/0
Copyright © 1989, American Society for Microbiology
Vol. 9, No. 8

The *DAL7* Promoter Consists of Multiple Elements That Cooperatively Mediate Regulation of the Gene's Expression

HYANG SOOK YOO† AND TERRANCE G. COOPER*

Department of Microbiology and Immunology, University of Tennessee, Memphis, Tennessee 38163

Received 1 February 1989/Accepted 25 April 1989

Expression of the allantoin system genes in *Saccharomyces cerevisiae* is induced by allophanate or its analog, oxalurate. This work provides evidence for the involvement of distinct types of *cis*-acting elements in the induction process. The first element was found to have the properties of an upstream activation sequence (UAS). This element was localized to a 16-base-pair (bp) DNA fragment containing a short 5-bp sequence that occurred repeatedly in the upstream region of *DAL7*. When present in two or more copies, the 16-bp fragment supported high-level β-galactosidase production in a *CYC1-lacZ* expression vector; there was, however, no response to the allantoin pathway inducer. The second element had the properties of a negatively acting element or upstream repression sequence (URS). This element was localized to a 16-bp DNA fragment containing an 8-bp sequence that was repeated four times in the upstream region of *DAL7*. A fragment containing the 8-bp repeated sequence placed adjacent to the UAS-containing fragment mediated inhibition of the ability of the UAS to support *lacZ* expression regardless of whether inducer was present. A third element, designated an upstream induction sequence (UIS), was required for response to inducer. The UIS was localized to a small DNA fragment containing an approximately 10-bp sequence that was repeated twice in the upstream region of *DAL7*. When a fragment containing the 10-bp repeated sequence was placed adjacent to these UAS and URS elements, the construction (UIS-UAS-URS) supported normal oxalurate-mediated induction of β-galactosidase synthesis. These data are consistent with the suggestion that multiple, *cis*-acting elements participate in the induction process.

The upstream activation sequences (UASs) of yeast cells, which have many features in common with mammalian cell enhancer sequences (3, 10, 18), have been convincingly shown to be a primary route of regulated gene expression in yeast cells (7, 13, 15). *GAL* gene expression has been correlated with binding of the *GAL4* gene product to a 17-base-pair (bp) region situated in the 5'-flanking regions of those genes (2, 9). Response of the *HIS* genes to activation by the *GCN4* gene product of the general amino acid regulation system is another case in which activation and regulation of gene expression were localized to a short nucleotide sequence, 5'-TGACTC-3' (7, 16). In both cases, regulation was hypothesized to occur through controlled production of the protein that activated transcription or through regulation of its availability in functional form.

Recent studies make it clear that multiple proteins may be associated with the functioning of even the smallest *cis*-acting elements (23), and multiple sites are increasingly being found to be required for the process of regulated gene expression (11, 38). The functions of these sites in the 5'-flanking region and the protein factors potentially associated with them are in many cases yet to be identified. This work was undertaken to determine whether multiple regulatory sites participate in control of the inducible allantoin system genes and, if so, to identify their potential roles in the induction process.

Allantoin degradation in *Saccharomyces cerevisiae* is accomplished by the products of a regulated set of genes. Four of the genes (*DAL4*, *DAL7*, *DUR1,2*, and *DUR3*) are highly inducible, two (*DAL1* and *DAL2*) are moderately inducible, and the remaining two (*DAL3* and *DAL5*) are expressed independently of inducer (5, 6, 8, 19, 24, 33, 37; R. Rai and T. G. Cooper, unpublished observations). Allophanate and oxalurate serve as native and gratuitous inducers, respectively, which trigger expression (29, 36). Transcripts of the allantoin system genes are markedly overproduced in *dal80* mutants (4, 8, 24, 37; Rai and Cooper, unpublished observations) and require, to a greater or lesser degree, a functional *DAL81* gene product for expression (8, 35, 37). The highly pleiotropic effects of the *dal80* and *dal81* mutations suggest that all eight pathway genes have one or more common steps in their regulation and therefore may contain common elements associated with this control in their 5'-flanking regions.

We have chosen to use the *DAL7* gene for our analysis of an inducible allantoin system promoter. The experiments presented here show that the *DAL7* 5'-flanking region contains repeated copies of three types of elements. Each element was localized to a small DNA fragment, and all were shown to be required for normal transcriptional induction. We further correlated the elements with potential roles in the induction process.

(Preliminary reports of this work have appeared elsewhere [Abstr. Annu. Meet. Genet. Soc. Am. 1986, p. 29; Yeast **2**:S37, 1986; Abstr. Genet. Soc. Am. Yeast Genet. Mol. Biol. Meet. 1987, p. 349].)

MATERIALS AND METHODS

Strains and culture conditions. The *S. cerevisiae* and *Escherichia coli* strains used were RH218 (*MAT***a** *trp1 CUP1 gal2 SUC2* Mal$^-$) and HB101 (*hsdR hsdM recA13 supE44 lacZ24 leuB proA2 thi-1* Sm^r), respectively. Yeast cells containing each of the plasmids used were grown in yeast carbon base (11.7 g/liter; Difco Laboratories). Proline (0.1%) was added as the sole nitrogen source. Oxalurate (66 mg/liter) was provided as inducer. Cells were grown to a cell

* Corresponding author.
† Present address: Genetic Engineering Center, Korea Advanced Institute of Science & Technology, Cheongryang, Seoul, Korea.

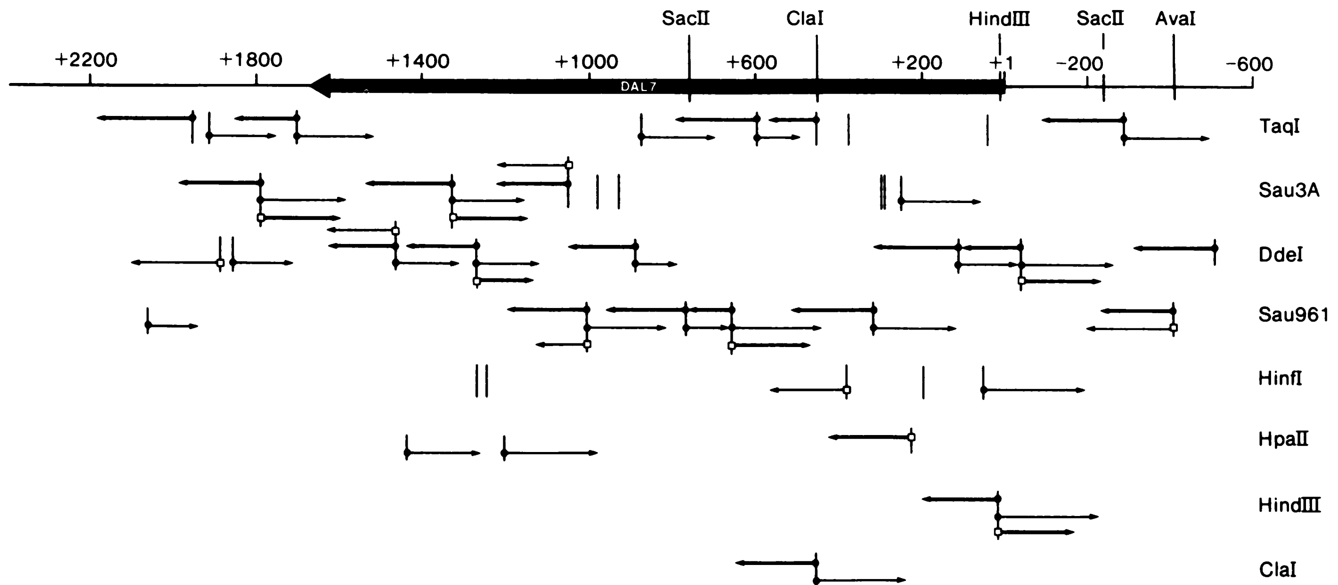

FIG. 1. DNA sequencing strategy for the *DAL7* gene. Fragments were labeled at the 5′ termini with polynucleotide kinase (●) or by filling in the 3′ termini with the Klenow fragment of DNA polymerase I (□). The length of the arrow indicates the number of nucleotide bases determined from that fragment. Restriction sites (|) used for labeling or digestion are indicated at the right. All positions are designated relative to the ATG of the coding sequence. By this strategy, 100% of the sequence for both strands of the DNA was determined, along with the sequence across every restriction site used for digestion or labeling. In most instances, the nucleotide sequences were read from more than one gel.

density of 40 to 45 Klett units and then harvested for enzyme assay. Transformation and plasmid manipulation procedures have been described elsewhere (30).

Nucleotide sequence analysis of the *DAL7* gene. The *DAL7* nucleotide sequence was determined by Maxam-Gilbert DNA sequence analysis as described earlier (21). A high-resolution restriction map of the *DAL7* gene and the strategy used to determine its nucleotide sequence are shown in Fig. 1. Both DNA strands were entirely sequenced, and every restriction site required for labeling was crossed. We primarily used 5′ end labeling with secondary digestion or strand separation to generate DNA fragments for sequence analysis. However, in several instances we labeled the DNA by filling in a recessed 3′ terminus, using the Klenow fragment of DNA polymerase I.

Determination of the 5′ termini of *DAL7*-specific RNA. The 5′ termini of *DAL7* transcripts were identified by procedures described earlier (32).

Plasmid constructions. The *CYC1-lacZ* expression vectors used (Fig. 2) were derivatives of plasmid pLG669Z (12). Oligonucleotides containing portions of the *DAL7* 5′-flanking sequences were synthesized with an Applied Biosystems oligonucleotide synthesizer. *Sal*I sites were synthesized on the ends of each oligonucleotide. Synthetic oligonucleotides were purified through 10% polyacrylamide-urea sequencing gels and then recovered by using the crush-soak method of Maxam and Gilbert (21). Purified, single-stranded oligonucleotides (1 to 5 μg) were dissolved in annealing buffer (20 mM Tris [pH 7.6], 1 mM $MgCl_2$). The two opposite strands of each desired fragment were combined, heated (95°C) for 10 min, and slowly cooled to room temperature. Annealed DNA fragments were precipitated by adding 1/10 volume of 3 M sodium acetate and 2 volumes of ethanol. A 1- to 2-μg sample of each annealed fragment preparation was phosphorylated by the polynucleotide kinase reaction (37°C for 1 h) as described by Maniatis et al. (20) and ligated into the *Xho*I site of plasmid pHY100. The structure of every construction used was verified by Maxam-Gilbert sequence analysis.

Construction of plasmids containing 5′ deletions in the *DAL7* promoter region. Plasmid pHY43 (Fig. 3), which contains sequences between positions −411 and +12 of the *DAL7* gene, was digested with endonuclease *Eco*RI and treated with nuclease BAL 31 for various lengths of time to generate a nested set of deletion fragments. After this treatment with Klenow fragment, *Bam*HI linkers (5′-CCCG GATCCGGG-3′) were ligated onto linear DNA fragments. After digestion of these DNA fragments with *Sac*I, the 2.2-kbp *Bam*HI-*Sac*I fragments were isolated and ligated into the 5.5-kbp *Bam*HI-*Sac*I fragment from plasmid pMC1790. To determine the deletion endpoints, each plasmid was digested with endonucleases *Bam*HI and *Sac*I, and the isolated deletion fragments were rendered radioactive by the polynucleotide kinase reaction and sequenced by Maxam-Gilbert procedures.

β-Galactosidase assay. β-Galactosidase activities in yeast transformants were determined by the method of Guarente and Mason (14). Units are those of Miller (22). Since all of the plasmids used in this work contain an autonomously replicating sequence, we took precautions to avoid problems that might result from varying copy numbers (25). All of the plasmids used in a given figure or table were transformed into the same sample host cells. Random transformants were isolated as soon as they were large enough to serve as inocula. These inocula were then grown up and assayed immediately; transformants were never subcultured or stored. Each experiment reported was repeated several times. Each repeat involved the generation and assay of new transformants. The absolute values of the assays varied somewhat, but the patterns of activity, including the subtle ones, observed from one construction to another were invariant. The data from repeated experiments generally varied less than 15%. Occasionally, spurious values were more than 30% out of line with results of repeated experi-

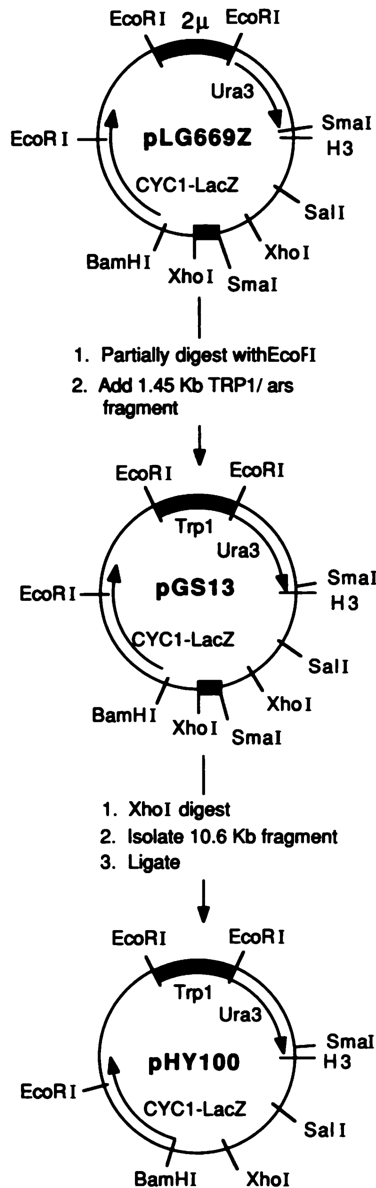

FIG. 2. Structures of vectors used. ars, Autonomously replicating sequence.

ments; in these cases, the spurious values could not be repeated.

RESULTS

Nucleotide sequence analysis of the *DAL7* gene. Our previous work localized the *DAL7* gene to a 2.4-kbp *ClaI* fragment derived from plasmid pTC7 (37). This fragment was used to generate the fine-structure restriction map (Fig. 1), using the Smith-Birstiel procedure (28). Using this map and the strategy depicted in Fig. 1, we determined the nucleotide sequence of a 2,360-bp region (Fig. 4). Analysis of the nucleotide sequence revealed an open reading frame of 1,662 bp that is predicted to encode a protein of 554 amino acids with a calculated monomer molecular weight of 62,796. Translation of the open reading frame predicts that the *DAL7* gene product is rather hydrophilic, has the overall structural characteristics of a globular protein, and contains three sites in the N-terminal 100 amino acids that are homologous with those commonly found to be glycosylated.

Mapping the *DAL7* transcript. We used Berk-Sharp S1 mapping procedures (1) and a 480-bp *HinfI-AvaI* fragment (Fig. 1), labeled at its 5′ terminus by the polynucleotide kinase reaction, to determine the 5′ terminus of the *DAL7* transcript (32). We observed one major protection fragment flanked by two minor fragments (Fig. 5). The major fragment was situated at position −25 relative to the ATG of the coding region. The S1 protection fragments we observed were strand specific and RNA and S1 nuclease dependent (Fig. 5).

Analysis of deletions in the *DAL7* 5′-flanking region. We began analysis of the *DAL7* regulatory region by constructing a set of nested deletion plasmids (see Materials and Methods). The deletion plasmids were transformed into wild-type strain RH218, and the transformants were assayed for β-galactosidase activity (Fig. 6). The first impression generated by the results of this analysis was that the patterns of enzyme activity and regulation were far more complex than previously observed with other genes, such as those of the *GAL* system. The starting plasmid (pHY43-1), which contained the *DAL7* 5′-flanking region to position −375, supported fully inducible β-galactosidase activity (Fig. 6). This finding indicated that sequences required for normal regulation of this gene were situated downstream of position −375. Plasmid pHY43-2 deleted all *DAL7* DNA upstream of position −339 and resulted in a modest but reproducible decrease in the uninduced levels of β-galactosidase production. The next deletion (plasmid pHY43-5) removed another 49 bp of DNA (to position −290) and resulted in a nearly 10-fold increase in the uninduced level of β-galactosidase expression. The appearance of a marked increase in the uninduced level of β-galactosidase production after deletion of a segment of DNA from a fusion plasmid is the expected result if the deletion removes a negatively acting site. This result and its interpretation have been previously documented for the *STE6* and *CAR1* genes (17, 31), among others. The next deletion (plasmid pHY43-6) removed all *DAL7* DNA upstream of position −256 and resulted in a decrease on β-galactosidase production regardless of whether inducer was present in the culture medium. The next deletion (plasmid pHY43-7) removed 20 bp of DNA and exhibited the most marked phenotypic change observed in the series of plasmids analyzed. Nearly all *lacZ* expression was lost, but the activity that remained was still inducible. We interpreted this result as meaning two things: that a central component of the remaining expression apparatus was located between positions −256 to −236, and that sequences downstream of position −236 retained sufficient information to support inducible gene expression. The next two deletions (plasmids pHY43-8 and pHY43-9) yielded the same results, indicating that sequences between positions −236 and −229 were not necessary for the induction process. Deletion of an additional 8 bp of DNA, to position −221 (plasmid pHY43-10), resulted in two effects. The remaining activity decreased by about half, and the activity that remained was unresponsive to the presence of inducer. The most straightforward interpretation of this result is that sequences between positions −229 and −221 were required for the response to inducer. Finally, deletion to position −205 resulted in elimination of all detectable activity, indicating that sequences between positions −221 and −205 contained sequences required for gene expression.

Although these patterns of expression and regulation appeared to be complex, the data pointed to the existence of

FIG. 3. Construction of plasmids used to generate a 5′ deletion map of the *DAL7* gene. ARS, Autonomously replicating sequence.

at least three types of regulatory sites: (i) a negatively acting site situated between positions −339 and −290, (ii) positively acting sites situated between positions −290 to −256 and −221 to −205, and (iii) a site that was somehow related to induction situated between positions −229 and −221. Alternatively, one could argue that the complex variation in β-galactosidase activity derived from variation in the assays we used. However, the patterns of activity observed among the plasmids were reproducible, which would not be expected if the variations were random. The two possible explanations for our data, one significant and the other trivial, were experimentally distinguishable. If the observed

FIG. 4. Nucleotide sequence of the *DAL7* gene and its 5′- and 3′-flanking regions. The inferred protein sequence is shown below the sense strand of the DNA sequence. The termini of S1 nuclease protection fragments (putative transcriptional start sites) are designated by arrows, and the putatively operational TATA sequence is underlined. The three sets of homologous sequences described in the text and in the legend to Fig. 6 are indicated with different symbols.

-400 -390 -380 -370 -360
CC CGA GAA GGA CTG TGT GGA AGA ACA CTA CAG CGA TGG CGA CGT TTG TAT TAT CAT CTA

-350 -340 -330 -320 -310 -300
ATT CTG CGA CTG TGG GCG AAG TAA ATA TGT AAA TGT TCA TGT ATT AGT AAT TTC ATA

-290 -280 -270 -260 -250 -240
GGT TTA CGA TTA TCG ACA TAA AAT CTC CGC TGA AAG TTG CGG TGC GAT AGA ATA CCG

-230 -220 -210 -200 -190
CGG ATT TTG GAA AAT TGC GTT TGC TTT TCT TAT CAC ATA AAT TGA ATT TCA TAA GTT

-180 -170 -160 -150 -140 -130
TTG GCT TAT CTT GGA TAA GGT ACA TCG TGA AAG TAT CAC AAA GAT TGG ATA GTT TCG

-120 -110 -100 -90 -80 -70
TCA ATT GGT GAA GCA ATA AGT ATA AAA TGG GGA CAT TTA CGC ATG TAG TCT TAA TTC

-60 -50 -40 -30 -20 -10
TTC ACT TTC TGT GCA GTT GAT ATC ACT TAG AGT ATG TGT CAT AGG CAC GGT AAA GAG

+1 +10 +20 +30 +40
CAC TGA ACA ATG GTG AAG ATA AGC TTG GAC AAC ACT GCT CTA TAC GCA GAC ATC GAC
Met Val Lys Ile Ser Leu Asp Asn Thr Ala Leu Tyr Ala Asp Ile Asp

+50 +60 +70 +80 +90 +100
ACG ACT CCT CAA TTT GAA CCT TCC AAA ACT ACT GTA GCT GAT ATT TTA ACA AAA GAT
Thr Thr Pro Gln Phe Glu Pro Ser Lys Thr Thr Val Ala Asp Ile Leu Thr Lys Asp

+110 +120 +130 +140 +150 +160
GCC TTA GAG TTC ATT GTT TTG CTG CAT AGA ACT TTC AAT TCA ACA CGG AAA CAG CTT
Ala Leu Glu Phe Ile Val Leu Leu His Arg Thr Phe Asn Ser Thr Arg Lys Gln Leu

+170 +180 +190 +200 +210
TTA GCC AAC AGA AGC AAT TTA CAA TCA AAG TTA GAC TCT GGT GAA TAC CGA TTT GAT
Leu Ala Asn Arg Ser Asn Leu Gln Ser Lys Leu Asp Ser Gly Glu Tyr Arg Phe Asp

+220 +230 +240 +250 +260 +270
TTC TTG CCG GAA ACC GAA CAA ATC AGG AAC GAT CCT ACA TGG CAA GGT GCT ATC CCA
Phe Leu Pro Glu Thr Glu Gln Ile Arg Asn Asp Pro Thr Trp Gln Gly Ala Ile Pro

+280 +290 +300 +310 +320 +330
GCC CCT GGT TTG ATC AAC AGA TCC AGC GAG ATT ACT GGG CCA CCA TTG AGA AAT ATG
Ala Pro Gly Leu Ile Asn Arg Ser Ser Glu Ile Thr Gly Pro Pro Leu Arg Asn Met

+340 +350 +360 +370 +380 +390
TTA GTC AAC GAT TTG AAT GCT GAA GTA ACG ACA TAT ATG ACT GAT TTC GAG GAC TCT
Leu Val Asn Asp Leu Asn Ala Glu Val Thr Thr Tyr Met Thr Asp Phe Glu Asp Ser

+400 +410 +420 +430 +440
TCA TCT CCA ACT TGG GAG AAC ATG ATT TAC GGA CAA GTT AAT CTT TAT GAT GCC ATT
Ser Ser Pro Thr Trp Glu Asn Met Ile Tyr Gly Gln Val Asn Leu Tyr Asp Ala Ile

+450 +460 +470 +480 +490 +500
AGA AAT CAA ATC GAT TTC AAG ACA CCA AGA AAG GAG TAC AGG TTG AAA GGT GAC ATT
Arg Asn Gln Ile Asp Phe Lys Thr Pro Arg Lys Glu Tyr Arg Leu Lys Gly Asp Ile

+510 +520 +530 +540 +550 +560
TCA AGA CTG CCC ACT TTA ATT GTC AGA CCT CGT GGC TGG CAC ATG GTG GAG AAG CAC
Ser Arg Leu Pro Thr Leu Ile Val Arg Pro Arg Gly Trp His Met Val Glu Lys His

```
        +570          +580          +590          +600          +610
          *             *             *             *             *
CTT TAC ATA GAT GAT GAA CCG ATT AGT GCT TCC ATC TTC GAT TTT GGT TTA TAT TTT
Leu Tyr Ile Asp Asp Glu Pro Ile Ser Ala Ser Ile Phe Asp Phe Gly Leu Tyr Phe

+620          +630          +640          +650          +660          +670
  *             *             *             *             *             *
TAC CAT AAC GCT AAA GAG TTA GTT AAA ATT GGT AAA GGA CCT TAT TTT TAC TTA CCA
Tyr His Asn Ala Lys Glu Leu Val Lys Ile Gly Lys Gly Pro Tyr Phe Tyr Leu Pro

    +680          +690          +700          +710          +720          +730
      *             *             *             *             *             *
AAG ATG GAG CAC CAT ATG GAG GTA AAA CTA TGG AAT GAC ATA TTC TGT GTT GCA CAA
Lys Met Glu His His Met Glu Val Lys Leu Trp Asn Asp Ile Phe Cys Val Ala Gln

        +740          +750          +760          +770          +780
          *             *             *             *             *
GAT TTT ATT GGA ATG CCC CGC GGT ACC ATT AGG GCC ACT GTT CTG ATT GAA ACT TTG
Asp Phe Ile Gly Met Pro Arg Gly Thr Ile Arg Ala Thr Val Leu Ile Glu Thr Leu

+790          +800          +810          +820          +830          +840
  *             *             *             *             *             *
CCA GCG GCC TTC CAA ATG GAG GAG ATT ATC TAT CAA ATA AGA GAA CAT TCA AGC GGT
Pro Ala Ala Phe Gln Met Glu Glu Ile Ile Tyr Gln Ile Arg Glu His Ser Ser Gly

    +850          +860          +870          +880          +890          +900
      *             *             *             *             *             *
TTG AAC TGT GGT CGT TGG GAC TAC ATA TTT TCG ACC ATT AAA AAA CTG AGA AAC TTG
Leu Asn Cys Gly Arg Trp Asp Tyr Ile Phe Ser Thr Ile Lys Lys Leu Arg Asn Leu

        +910          +920          +930          +940          +950          +960
          *             *             *             *             *             *
AAT GAA CAC GTT TTG CCA AAT AGG GAT CTA GTG ACT ATG ACT TCA CCT TTT ATG GAT
Asn Glu His Val Leu Pro Asn Arg Asp Leu Val Thr Met Thr Ser Pro Phe Met Asp

            +970          +980          +990          +1000         +1010
              *             *             *             *             *
GCT TAT GTG AAA AGA TTG ATC AAT ACA TGT CAC CGT AGA GGG GTC CAT GCG ATG GGT
Ala Tyr Val Lys Arg Leu Ile Asn Thr Cys His Arg Arg Gly Val His Ala Met Gly

+1020         +1030         +1040         +1050         +1060         +1070
  *             *             *             *             *             *
GGT ATG GCT GCC CAA ATC CCC ATA AAA GAT GAT CCA AAG GCT AAT GAA GCT GCA ATG
Gly Met Ala Ala Gln Ile Pro Ile Lys Asp Asp Pro Lys Ala Asn Glu Ala Ala Met

    +1080         +1090         +1100         +1110         +1120         +1130
      *             *             *             *             *             *
AAC AAA GTT CGT AAT GAC AAA ATT AGA GAA ATG AAG AAT GGG CAT GAT GGG TCA TGG
Asn Lys Val Arg Asn Asp Lys Ile Arg Glu Met Lys Asn Gly His Asp Gly Ser Trp

        +1140         +1150         +1160         +1170         +1180
          *             *             *             *             *
GTA GCA CAC CCA GCA TTG GCA CCG ATT TGT AAT GAA GTT TTC AGT AAC ATG GGT ACA
Val Ala His Pro Ala Leu Ala Pro Ile Cys Asn Glu Val Phe Ser Asn Met Gly Thr

+1190         +1200         +1210         +1220         +1230         +1240
  *             *             *             *             *             *
GCA AAT CAA ATA TAT TTT GTC CCG GAT GTA CAT GTT ACA TCA TCT GAT TTA TTG AAT
Ala Asn Gln Ile Tyr Phe Val Pro Asp Val His Val Thr Ser Ser Asp Leu Leu Asn

    +1250         +1260         +1270         +1280         +1290         +1300
      *             *             *             *             *             *
ACG AAG ATT CAA GAT GCT CAA GTC ACT ACT GAG GGA ATC AGA GTA AAC TTG GAT ATT
Thr Lys Ile Gln Asp Ala Gln Val Thr Thr Glu Gly Ile Arg Val Asn Leu Asp Ile

        +1310         +1320         +1330         +1340         +1350
          *             *             *             *             *
GGC CTA CAA TAT ATG GAG GCT TGG TTA AGG GGA TCT GGT TGT GTC CCA ATT AAT CAT
Gly Leu Gln Tyr Met Glu Ala Trp Leu Arg Gly Ser Gly Cys Val Pro Ile Asn His

+1360         +1370         +1380         +1390         +1400         +1410
  *             *             *             *             *             *
TTG ATG GAA GAT GCC GCT ACT GCG GAA GTA TCA CGT TGT CAA TTG TAC CAG TGG GTT
Leu Met Glu Asp Ala Ala Thr Ala Glu Val Ser Arg Cys Gln Leu Tyr Gln Trp Val

    +1420         +1430         +1440         +1450         +1460         +1470
      *             *             *             *             *             *
AAA CAT GGT GTT GTC TTA AGT GAT ACC GGT GAC AAA GTA ACT CCA GAA TTG ACC GCT
Lys His Gly Val Val Leu Ser Asp Thr Gly Asp Lys Val Thr Pro Glu Leu Thr Ala

        +1480         +1490         +1500         +1510         +1520         +1530
          *             *             *             *             *             *
AAG ATA TTA AAT GAA GAG ACT GCA AAA TTG GCT TCA GCA AGT CCG CTG GGT GAA AAG
Lys Ile Leu Asn Glu Glu Thr Ala Lys Leu Ala Ser Ala Ser Pro Leu Gly Glu Lys
```

FIG. 4—*Continued*

```
         +1540     +1550     +1560      +1570     +1580
           *         *         *          *         *
AAC AAG TTT GCC TTG GCA GCC AAG TAT TTT TTG CCT GAA GTC ACT GGT AAA ATC TTT
Asn Lys Phe Ala Leu Ala Ala Lys Tyr Phe Leu Pro Glu Val Thr Gly Lys Ile Phe

+1590      +1600     +1610     +1620      +1630     +1640
  *          *         *         *          *         *
AGC GAC TTC TTG ACC ACT TTA TTG TAT GAT GAA ATT ATT AAG CCA AGT GCC AAA CCA
Ser Asp Phe Leu Thr Thr Leu Leu Tyr Asp Glu Ile Ile Lys Pro Ser Ala Lys Pro

  +1650      +1660     +1670     +1680      +1690     +1700
    *          *         *         *          *         *
GTT GAC TTA AGT AAA TTA TAG AAT GTA TAC GTA CAT AAC CTG ACG AAT ATT CGA AGA
Val Asp Leu Ser Lys Leu ---

       +1710      +1720     +1730     +1740      +1750
         *          *         *         *          *
ATT TTA GGC GCT GTT TTA ACG CAT AAT CCA ATC AAA TTT GAG CTG ATA AGG CCT GAT

+1760      +1770      +1780     +1790     +1800      +1810
  *          *          *         *         *          *
AAG GTA TGA CAA ATT GAA CTC ATA TTT TTT CTC CAT GCC AGA TCA TAT TTA AGC CTT

  +1820      +1830      +1840     +1850     +1860      +1870
    *          *          *         *         *          *
TGA AGC GAG AAT ATG TAA GGA AAC TGA ATT ACT AAT TCT TAC CTC AGG AAA ATC AAA

       +1880     +1890      +1900     +1910     +1920
         *         *          *         *         *
AGA ACA AAG AAA AAC TAG CTA AGC AAA TTA TCG AAC GAT GGA AAC AAG AAT ACT TGT

+1930     +1940     +1950
  *         *         *
TGT GAA TCC TAA TAG TTC GAA
```

FIG. 4—*Continued*

levels of *lacZ* expression supported by this set of plasmids derived from random variation in plasmid numbers or other random technical problems, they should not lead to predictable results upon further testing. If, on the other hand, the variations in β-galactosidase levels were physiologically significant, we reasoned that it should be possible to isolate the predicted elements and demonstrate their characteristics alone and in combination. Therefore, our next objective was to isolate and characterize each of the three elements suggested to exist by the deletion analysis.

Meeting this objective was aided significantly by the observation that each of the three regions indicated by deletion analysis to potentially contain a regulatory sequence also contained a short sequence that was repeated throughout the *DAL7* 5′-flanking region. It was also of great advantage that one small region upstream of the *DAL7* TATA box contained all three of these potentially important sequences (positions −233 to −183). This fact permitted the first test of our working hypothesis. If the three types of repeated sequences identified by deletion analysis are all that is required for regulated gene expression, then the 51-bp DNA fragment containing them should support normal induction when inserted into the heterologous *CYC1-lacZ* expression vector (plasmid pHY100). By itself, this vector would not support *lacZ* expression (Fig. 7). To test this prediction, an oligonucleotide covering positions −233 to −183 was synthesized, and two copies of it were inserted in tandem into the *Sal*I site of plasmid pHY100. The purpose of inserting two copies of the sequence in tandem was predicated on the fact that one of the sequences contained in the oligonucleotide was identical to one that had been previously shown to function as the UAS of the *DAL5* gene, but to do so only when present in two or more copies (25). Similar results have also been reported for the *SUC2* and *CUP1* genes (26, 34). The resultant plasmid, pHY174, was transformed into strain RH218, and the strain was assayed for β-galactosidase production. The sequences cloned into the expression vector were able to support β-galactosidase production, which increased 18-fold when the allantoin system gratuitous inducer (oxalurate) was added to the culture medium (Fig. 7). This result argued that the 51-bp fragment, selected solely on the basis of possessing the three repeated sequences, was sufficient to support inducible gene expression. Our next objective was to determine whether this 50-bp fragment could be dissected into subfragments, each containing one of the repeated sequences identified by the deletion analysis as potentially playing a role in induction.

Identification of the *DAL7* UAS. The first DNA fragment to be isolated from the 51-bp sequence described above behaved as though it contained a UAS. We first focused our attention on this DNA fragment because its sequence was nearly identical to one that had been previously shown to be the UAS of the *DAL5* gene (25). The pertinent information derived from the *DAL5* experiments is as follows.

The *DAL5* gene was shown to contain six copies of the pentanucleotide 5′-GATAA-3′ or its reverse complement 5″-TTATC-3″. Deletions that successively removed copies of these pentanucleotides resulted in a progressive decrease in *DAL5* gene expression. This result implicated the sequence 5′-GATAA-3′ as part of the *DAL5* UAS. Further support for the suggestion that DNA fragments containing the pentanucleotide were capable of mediating transcriptional activation derived from the observation that two or more copies of a synthetic 11-bp oligonucleotide containing the sequence 5′-GATAA-3′ supported high-level β-galactosidase produc-

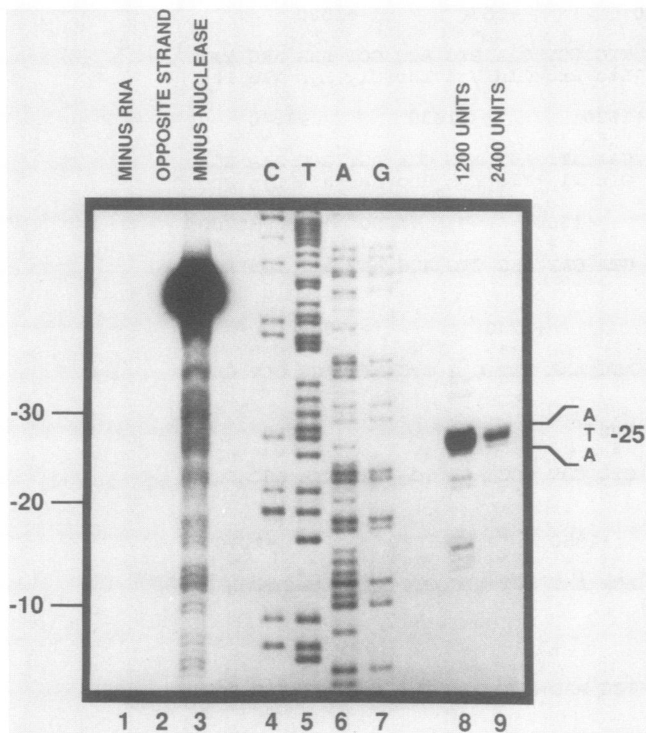

FIG. 5. Determination of the 5′ termini for *DAL7* transcripts by means of S1 nuclease protection. Poly(A)$^+$ RNA (8 μg) was hybridized with the labeled probe described in the text. The hybridization mixture was then treated with 1,200 (lane 8) or 2,400 (lane 9) U of S1 nuclease at 23°C for 1 h. The protected fragments were resolved on a 10% polyacrylamide gel together with products of the four Maxam-Gilbert sequencing reactions (lanes 4 to 7). The results of control reactions are shown in lanes 1 to 3.

tion when inserted into the heterologous *CYC1-lacZ* expression vector (plasmid pHY100). A single copy of the fragment was nonfunctional. Finally, the pentanucleotide was also shown to be situated beneath the footprint of one or more DNA-binding proteins (25).

The same pentanucleotide (5′-GATAA-3′ or its reverse complement) found in the *DAL5* upstream region was also found five times in the 5′-flanking region of the inducible *DAL7* gene (Fig. 4). The related sequences 5′-GATAG-3′ and 5′-GATGA-3′ appeared three more times. The existence of these repeated sequences in the upstream region of *DAL7* and their presence in the 51-bp sequence described above raised the possibility that they might function as the UAS for the *DAL7* gene as well. The ability of such a *DAL7*-derived DNA fragment to support heterologous gene expression was directly demonstrated by the data shown in Fig. 7. The *CYC1-lacZ* expression vector (plasmid pHY100; Fig. 3), lacking an insert at the *Xho*I site, supported 4 to 5 U of β-galactosidase production after transformation into wild-type yeast strain RH218; this value established the basal level of expression supported by vector sequences. When one copy of a 16-bp oligonucleotide, derived from positions −218 to −203 of the *DAL7* 5′-flanking sequence, was cloned into the *Xho*I site (plasmid pHY126; data not shown) and the plasmid was transformed into strain RH218, only 8 U of β-galactosidase was produced regardless of whether inducer (oxalurate) was present in the culture medium. In other words, one copy of the 16-bp sequence was unable to support gene expression, as previously shown for *DAL5* (25). When two copies of the 16-mer were tandemly cloned into the *Xho*I site, high-level β-galactosidase production was observed (plasmid pHY129) in both the absence (504 U of activity) and presence (440 U of activity) of inducer (Fig. 7). As a result of this observation, two copies of all subsequent constructions were cloned into the heterologous expression vector. These data suggested that sequences between positions −218 and −203 contained a yeast UAS that supported oxalurate-independent gene expression. These sequences contained the 5′-GATAA-3′ sequence, as did the DNA fragment between positions −256 and −236. The latter fact is important because deletions of these 20 bp resulted in a dramatic decrease in *DAL7* gene expression (Fig. 6).

Identification of a negatively acting site in *DAL7*. A DNA fragment with the characteristics of a negatively acting site or upstream repression sequence (URS) was identified by repeating the experiment described above but using an oligonucleotide (plasmid pHY156) that extended 12 bp further in the 3′ direction than did sequences in plasmid pHY129; the insert covered positions −218 to −192. The 12-bp fragment added onto the 3′ end of the insert carried by plasmid pHY129 contained a sequence very similar to the one also found between positions −298 and −291. These positions were included among those identified by deletion analysis as potentially containing a negatively acting site (plasmids pHY43-2 and pHY43-5; Fig. 6). Addition of the 12 bp of DNA between positions −203 and −192 resulted in complete loss of ability to express the *lacZ* gene (Fig. 7). The presence of the additional 12 bp inhibited the ability of sequences between positions −218 and −203 to support gene expression as they had in plasmid pHY129. This result suggested that a URS was situated between positions −203 and −192. The sequence contained between these positions was also repeated between positions −188 to −181 and −278 to −270 as well as in the two locations cited above. The same result was observed when this experiment was repeated using a fragment that covered positions −218 to −183 and hence contained two copies of the putative URS. A transformant containing the expression vector with a DNA fragment covering positions −203 to −183 (two copies of the URS alone) inserted into the *Xho*I site (plasmid pHY157) also failed to support *lacZ* expression (Fig. 7).

To test directly whether DNA sequences between positions −203 and −183 carried a URS with characteristics like those reported in the *CAR1* and *STE6* genes (17, 31), we cloned a DNA fragment containing these sequences into each of the two *Xho*I sites of plasmid pGS13 (Fig. 8). Plasmid pGS13 is a *CYC1-lacZ* fusion plasmid that contains an intact, wild-type *CYC1* UAS in its normal position between the two *Xho*I sites (Fig. 3). If sequences between positions −203 and −183 contained a URS, they would be expected to inhibit operation of the *CYC1* UAS, as has been previously shown for the URSs of *CAR1* (31) and *STE6* (17). Plasmid pHY230, which contained sequences between positions −203 and −183, repressed *lacZ* expression threefold (Fig. 8). Comparison of the results obtained with plasmids pHY230 and pHY234 indicated that inhibition of UAS function occurred only when the DNA fragment covering positions −203 to −183 was placed 3′ of the *CYC1* UAS, not 5′ of it. Ineffectiveness at the 5′ position may derive from the distance (435 bp) between the two *Xho*I sites.

Identification of a sequence associated with *DAL7* induction. The third type of element, one that we have designated the upstream induction sequence (UIS), was predicted to exist from the deletion experiment depicted in Fig. 6 and from experiments demonstrating that repression of gene activa-

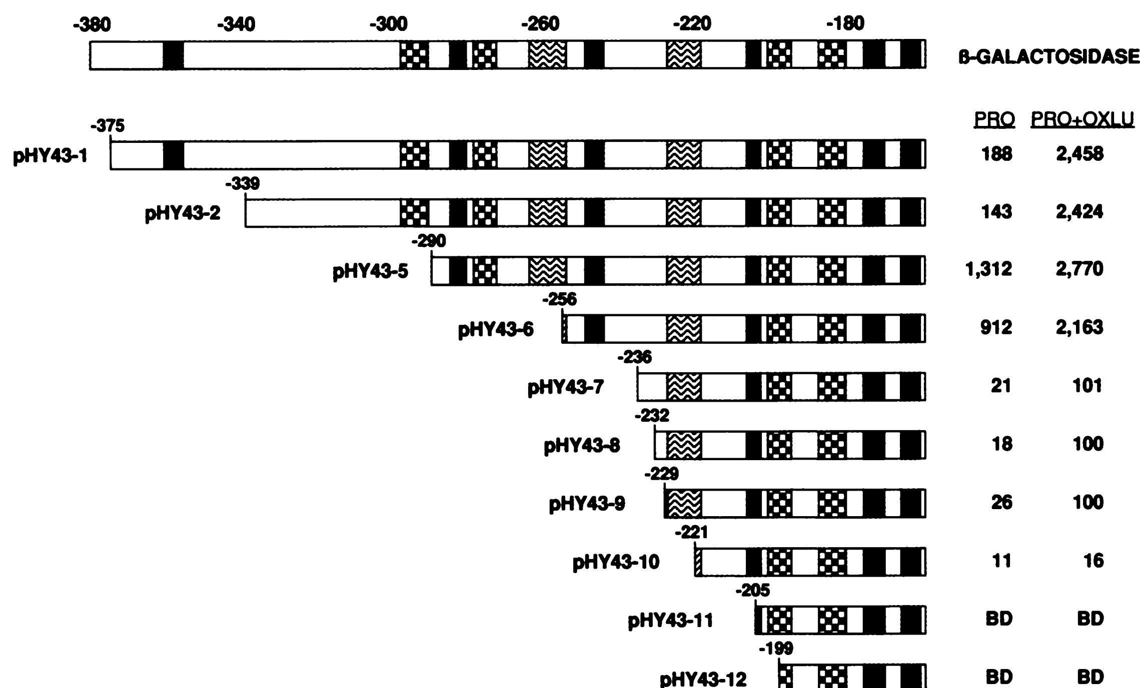

FIG. 6. Deletion analysis of the 5′-flanking region of the *DAL7* gene. The deletions were constructed as described in Materials and Methods and the legend to Fig. 3. The plasmids were then transformed into strain RH218. Transformants were grown and β-galactosidase was assayed as described in Materials and Methods. The deletion endpoints of the plasmids were −375, −339, −290, −256, −236, −232, −229, −221, −205, and −199 for plasmids pHY43-1, pHY43-2, pHY43-5, pHY43-6, pHY43-7, pHY43-8, pHY43-9, pHY43-10, pHY43-11, and pHY43-12, respectively. The filled, checkered, and lined areas indicate the positions of UASs, URSs, and UISs, respectively. The limits of the regulatory sequences have been set in this and other figures on the basis of homology. Coordinates are designated relative to the ATG of the coding sequence.

tion mediated by the *CAR1* URS was prevented during induction only as long as the URS was situated upstream of the *CAR1* gene (31). In other words, the ability of the *CAR1* URS to function was observed to be *cis*-dominantly regulated, and an upstream induction sequence was hypothesized to mediate that regulation (31).

Data presented in Fig. 7 demonstrate the existence of two *cis*-acting elements, one that appears to be involved in the activation of gene expression and a second that seems to mediate inhibition of that activation. Neither of these sequences alone or in combination, however, could be demonstrated to mediate the 10- to 15-fold induction observed by Northern (RNA) analysis of the wild-type *DAL7* gene (37), with *DAL7-lacZ* fusions (Fig. 6), or with the 51-bp sequence in plasmid pHY174 (Fig. 7). The sequence that appeared to mediate induction was localized by comparing the patterns of β-galactosidase production supported by plasmids pHY135 and pHY174 (Fig. 7). These plasmids differed by only 11 bp. As already mentioned, plasmid pHY135 contained sequences between positions −218 and −183 inserted into the *Xho*I site of expression vector pHY100. The insert of plasmid pHY174, on the other hand, extended upstream to position −228. It was deletion of sequences between positions −229 and −221 (plasmid pHY43-10; Fig. 6) that resulted in loss of response to inducer. Plasmid pHY135 supported only 9 U of β-galactosidase activity regardless of whether oxalurate was present in the culture medium (Fig. 7). In sharp contrast to this result, plasmid pHY174, containing the additional 11 bp on its 5′ end, supported β-galactosidase production that was fully inducible. In other words, the response to induction depended on sequences between positions −218 and −228.

We concluded above that plasmid pHY135 failed to support β-galactosidase production because the URS it contained inhibited operation of a UAS situated between positions −209 and −203. By this reasoning, plasmid pHY174 was able to support inducible β-galactosidase production because the element contained in the 11-bp sequence between positions −228 and −218 was able to mediate, in an as yet unknown way, inhibition of URS-mediated repression of the UAS. An alternative possibility was that sequences between positions −228 and −218 contained an inducible UAS. To assess this possibility, we cloned a fragment containing these sequences into the expression vector pHY100. This plasmid (pHY144) was not able to support *lacZ* expression under any condition (Fig. 7).

If the entire set of three regulatory elements is required for normal induction, it is appropriate to determine the phenotype of a construction consisting of only the UIS and UAS elements. Such a construction (plasmid pHY162) was prepared and assayed. The uninduced level of β-galactosidase production supported by a plasmid containing both elements was higher (616 U) than that observed when only the UAS was present (504 U). Plasmid pHY162, however, exhibited only a 2-fold-greater response to inducer, compared with the 18-fold increase observed with plasmid pHY174, which contained all three elements. The modest increase in induction did not result from failure to stimulate expression when oxalurate was present; induction was nearly threefold higher in plasmid pHY162 than in plasmid pHY129. Rather, the induction supported by plasmid pHY162 derived from the fact that basal-level expression was so high (616 U). This pointed to the role of the URS in maintaining a low level of expression under uninduced conditions.

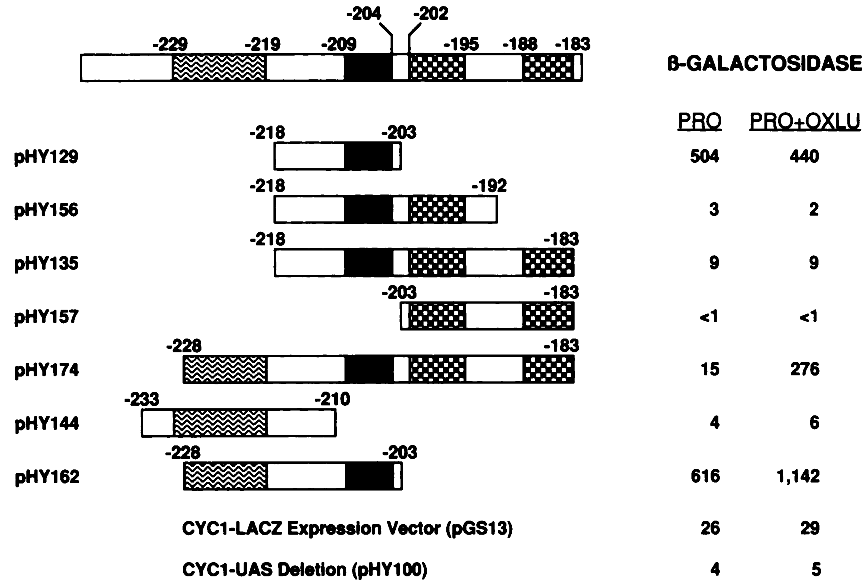

FIG. 7. *lacZ* expression mediated by synthetic oligonucleotides derived from the upstream region of the *DAL7* gene cloned into a *CYC1-lacZ* expression vector (pHY100; Fig. 3). Positions are given relative to the *DAL7* coding region. The structure of each plasmid was verified by sequence analysis. DNA sequences covering the positions indicated, with a *Sal*I site added to both ends, were synthesized as described in Materials and Methods. Purified, double-stranded fragments (two tandem copies) were then cloned into the *Xho*I site of *CYC1-lacZ* expression vector pHY100. After transformation of yeast strain RH218 with these plasmids, β-galactosidase activity was assayed by using cells grown in yeast carbon base medium containing proline as the sole nitrogen source. Yeast cultures were grown in either the absence (PRO) or presence (PRO + OXLU) of inducer. Regulatory sequences are indicated as in Fig. 6.

The preceding observations provide evidence that the UIS element somehow mediates inactivation of URS-mediated repression of transcription in the presence of inducer, thereby accomplishing induction when all three sequences are present. The UIS element also seems, however, to enhance transcriptional activation when placed adjacent to the UAS. This effect can be observed by comparing the level of β-galactosidase expression supported by the UAS element alone (504 U; plasmid pHY129 in Fig. 7) with those observed when the UIS and UAS elements are present together (616 U; plasmid pHY162 in Fig. 7). Here, the degree of UIS-mediated enhancement of UAS-mediated gene expression was only 20%, which is a far from compelling increase. However, when this experiment was repeated using a Σ1278b-related strain instead of strain RH218 as the transformation recipient, a 2.4-fold UIS-mediated enhancement of gene expression was observed between plasmids pHY129 and pHY162 (P. Bricmont and T. G. Cooper, unpublished observations).

Correlation of phenotypes observed with deletions in the *DAL7* upstream region and the nature of sequences lost in each deletion. Data presented above suggest that the *DAL7* upstream region contains multiple *cis*-acting elements that can be grouped on the basis of function. If it is assumed that the elements consist of DNA sequences that are homologous within each functional group, these homologies should be

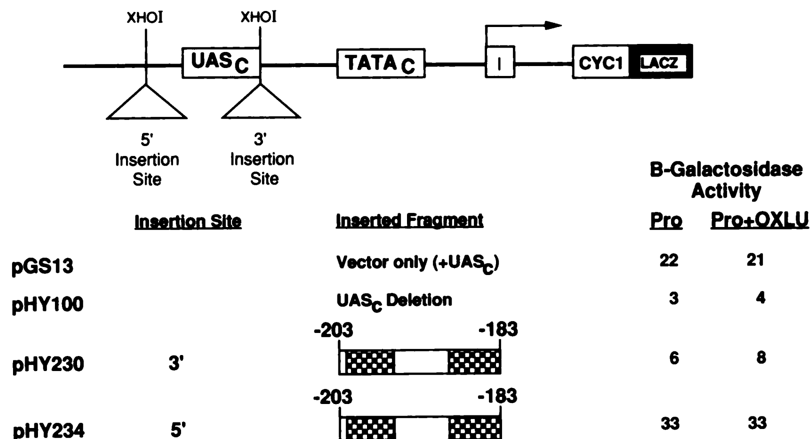

FIG. 8. Repression of the *CYC1* UAS by synthetic oligonucleotide fragments derived from the upstream region of the *DAL7* gene. Plasmids containing one copy of the fragment were transformed into strain RH218. Transformants were grown and assayed for β-galactosidase activity as described in the legend to Fig. 6. Checkered areas indicate locations of the URSs that have been designated on the basis of sequence homologies.

identifiable. Furthermore, we should be able to correlate the phenotype generated by each BAL 31 deletion with the homologous sequence(s) and nature of the putative element(s) that is removed. In such an analysis, 12 homologous sequences were observed (Fig. 4 and 6). It should be emphasized that this correlation is based solely on observed homology between sequences in different portions of the *DAL7* upstream region. There is not, at present, any mutational data that would support either the precise identity of the putative element sequences or the significance of the homologies noted.

Plasmid pHY43-1, used as a starting point for the analysis, supported a 13-fold increase in *lacZ* expression upon induction. Deletion pHY43-2 removed a sequence that is homologous to one that we have suggested to serve as a UAS. This correlated with the observed decrease in basal level activity by 25%. It should be noted, however, that deletion of the individual UAS elements did not show equivalent loss of function. This is the same phenotype that we reported for deletions that removed UASs from the *DAL5* upstream region (25). The next deletion (plasmid pHY43-5) removed a sequence homologous to one that we have suggested serves as a URS. This correlated well with a ninefold increase in the basal level of activity. The same phenotype was observed for deletion of the *CAR1* URS (31). Note the similarity of this result and those observed with plasmids pHY174 and pHY162 in Fig. 7. The next deletion, plasmid pHY43-6, resulted in removal of sequences homologous to three putative elements, a UAS, a URS, and a UIS. This deletion resulted in a decrease in both induced and uninduced *lacZ* expression and points out one of the limitations of this type of analysis. If more than a single element is lost at a time, it is not possible to unambiguously establish cause-effect relationships because the experiment contains multiple variables. The same would be true if one element mediated more than a single effect. Removal of sequences homologous to the UAS, as occurred in deletion plasmid pHY43-7, correlated well with a dramatic loss of β-galactosidase production. However, the activity that remained was still fivefold inducible. This, we suggest, is consistent with the fact that at least one copy of the UIS and URS elements and two copies of the UAS element remained. The next two deletions did not remove sequences that were homologous to any of the elements discussed above and, as expected, yielded the same phenotypes as did plasmid pHY43-7. Deletion of the final sequence homologous to one that we suggest is a UIS in this gene (plasmid pHY43-10) resulted in two effects. First, the basal level of β-galactosidase activity dropped in a manner similar to that noted for plasmids pHY43-5 and pHY43-6 (Fig. 6) and for plasmids pHY162 and pHY129 (Fig. 7). Second, β-galactosidase production was no longer inducible. The low level of inducer-independent synthesis that was supported by deletion plasmid pHY43-10 was similar to that observed for plasmids pHY156 and pHY135 in Fig. 7. Again, the correlation between removal of the multiply represented sequences and the phenotypes observed was strong. Finally, removal of one more sequence that was homologous to one that we have suggested serves as a UAS, as done by deletion pHY43-11 (Fig. 6), resulted in a loss of detectable enzyme activity, suggesting that the UAS-homologous sequences present downstream of the two URS-homologous sequences were insufficient to overcome repressive effects of the latter. Alternatively, the UAS-homologous sequences may have failed to function as UASs because of their positions.

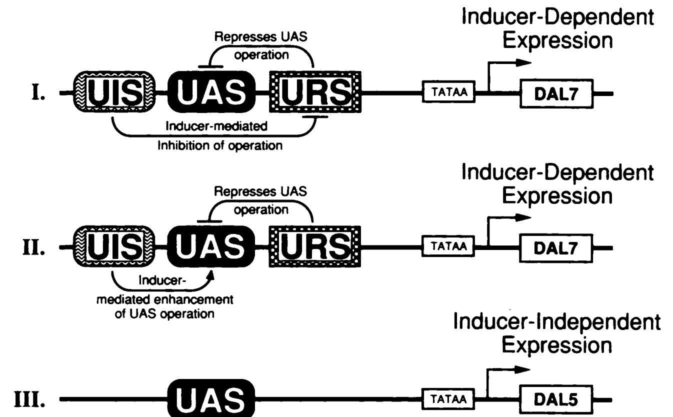

FIG. 9. Working models describing the promoter organization of the constitutive *DAL5* and inducible *DAL7* genes in *S. cerevisiae*. The models do not imply a specific order to the regulatory elements. Details are discussed in the text. Regulatory sequences are indicated as in Fig. 6.

DISCUSSION

The results of this analysis and work with the *CAR1* and *DAL5* genes (25, 31) generate at least two working models as possible explanations of how the *DAL7* gene is induced (Fig. 9). According to both models, transcriptional regulation of the *DAL7* gene is mediated by three functionally different *cis*-acting elements. The first model (I in Fig. 9) suggests that one *cis*-acting element, designated UAS, serves as the target site for a positively acting factor that activates transcription in an inducer-independent fashion. The model proposes the existence of a second region, designated URS, that serves as the target site for a negatively acting factor that represses operation of the activation factor(s) binding to the UAS. This could conceivably be accomplished by the URS-associated factor(s) interacting with the UAS-associated factor(s) as indicated. It is equally possible, however, that the URS element functions in some other way. A third region, designated UIS, is suggested to serve as the target site for a negatively acting factor that inhibits operation of the factor targeted to the URS. Production or functioning of the factor associated with the UIS is hypothesized to require the presence of inducer. By this model, induction results from the sum of two negative events rather than from a single positive one.

This explanation of the results was derived from the patterns of gene expression observed to be supported by the small DNA fragments described in Fig. 7. It assumes that each of the three sequences identified by the ability to alter the expression characteristics of the reporter gene are the target sites for transcription factors. If the assumption is valid, each of the three regions should be found to footprint in a DNA-binding assay. A first step in testing the validity of the assumption has already been accomplished for one of the sites with the recent demonstration of a footprint covering the DNA region containing a UAS like the one described in this work (25). Experiments are in progress to determine whether the remaining two regions are similarly protected in DNA footprinting assays.

Although the above explanation of our experimental results is quite consistent with the observed patterns of gene expression, it fails to account for the observed UIS-mediated enhancement of UAS function when only these two elements are present in the expression vector (plasmid

pHY162; Fig. 7). One way of accounting for the observation is to suggest that the UIS-associated factor(s) might perform two functions: (i) enhancement of UAS-mediated activation of gene expression and (ii) inhibition of URS-mediated repression of this activation. It is just as conceivable that the UIS element mediates only a single function, enhancement of UAS-mediated transcriptional activation; such a synergistic interaction has been hypothesized to occur in mammalian cells (27). By this interpretation of the data, the UIS-associated factor would mediate induction by enhancing UAS-mediated transcriptional activation to a point where its inhibition by the URS-associated element is overcome (model II in Fig. 9). What distinguishes the two models is whether the UIS-associated factor interacts with the UAS factor or the URS factor.

The data presented in this work correlate nucleotide sequences that are repeated in the 5′-flanking region of the *DAL7* gene with formal roles they seem to fulfill in the induction process. It is important to emphasize that the work focuses on the roles of the elements in the induction process, i.e., bringing about gene expression in response to inducer. Our experiments do not address the roles of these sites or the putative transcription factors they might bind in the process of transcription itself. The process of induction probably involves the inducer-dependent assembly of a preinitiation complex that is similar to those reported for more thoroughly studied genes. At least some of the factors for which we have ascribed roles in the induction process are probably in reality general transcription factors that play the same roles in *DAL7* transcription that they do elsewhere. At this point we have no information on what those roles might be. In this regard, the allantoin system genes need have only one specific factor, the one that recognizes the system inducer. The data we have presented suggest that the site associated with that factor is probably the UIS.

It is pertinent to distinguish the mechanistic differences between inducer-independent *DAL5* gene expression and inducer-dependent *DAL7* expression. The data presented here and those recently reported by Rai et al. (25) suggest that the difference between the responses of two genes to inducer derives from the element compositions in their 5′-flanking regions. The *DAL5* gene appears to possess only the UAS (Fig. 9), which correlates well with the fact that it also fails to respond to the allantoin system inducer (24, 25). The *DAL7* gene, on the other hand, has one element in common with the *DAL5* gene and two more that do not appear to be contained in the upstream region of *DAL5*.

This work demonstrates the complexity of the *DAL7* promoter. It appears to consist of multiple copies of three distinct types of regulatory sites. These sites are in addition to the TATA sequence and transcriptional start sites. As far as we can determine at present, most if not all of the *DAL7* elements identified by sequence homology appear to function. The strongest evidence in support of this view emanates from the deletion analysis (Fig. 6). The structure of the *DAL7* promoter suggests that there may be little unused DNA in the 5′-flanking region of this gene, since the three types of repeated elements described here appear to be quite closely spaced. If the sites identified above are transcription factor-binding sites, as has already been shown for the UAS site (25), the protein complexes formed at transcriptional initiation may be far larger and more complex than we have previously thought.

ACKNOWLEDGMENTS

We thank Roberta A. Sumrada for synthesizing all of the oligonucleotides used in this work. Thanks are also due those members of the University of Tennessee research group that read the manuscript and offered suggestions for improvement.

This work was supported by Public Health Service grants GM-19386, GM-35642, and GM-35536 from the National Institute of General Medical Sciences.

LITERATURE CITED

1. **Berk, A. J., and P. A. Sharp.** 1978. Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. USA **75**:1274–1278.
2. **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.
3. **Chambon, P., A. Dierich, M. Gaub, S. Jakowley, J. Jongstra, A. Krust, J. Lepennec, P. Oudet, and T. Reudelhuber.** 1984. Promoter elements of genes coding for proteins and modulation of transcription by estrogens and progesterone. Recent Prog. Hormone Res. **40**:1–39.
4. **Chisholm, G., and T. G. Cooper.** 1982. Isolation and characterization of mutants that produce the allantoin-degrading enzymes constitutively in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **2**:1088–1095.
5. **Cooper, T. G., and R. P. Lawther.** 1973. Induction of the allantoin degradative enzymes in *Saccharomyces cerevisiae* by the last intermediate of the pathway. Proc. Natl. Acad. Sci. USA **70**:2340–2344.
6. **Cooper, T. G., and R. Sumrada.** 1975. Urea transport in *Saccharomyces cerevisiae*. J. Bacteriol. **121**:571–576.
7. **Donahue, T. F., R. S. Daves, G. Lucchini, and G. R. Fink.** 1983. A short nucleotide sequence required for regulation of *HIS4* by the general control system of yeast. Cell **32**:89–98.
8. **Genbauffe, F. S., and T. G. Cooper.** 1986. Induction and repression of the urea amidolyase gene in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **6**:3954–3964.
9. **Giniger, E., S. M. Varnum, and M. Ptashne.** 1985. Specific DNA binding of *GAL4*, a positive regulatory protein of yeast. Cell **40**:767–774.
10. **Gluzman, Y., and T. Shenk.** 1983. Enhancers and eukaryotic gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Goodbourn, H., H. Burstein, and T. Maniatis.** 1986. The human β-interferon gene enhancer is under negative control. Cell **45**:601–610.
12. **Guarente, L.** 1983. Yeast promoters and *lacZ* fusions designed to study expression of cloned genes in yeast. Methods Enzymol. **101**:181–191.
13. **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**:504–511.
14. **Guarente, L., and T. Mason.** 1983. Heme regulatory transcription of the *CYC1* gene of *S. cerevisiae* via an upstream activation site. Cell **32**:1279–1286.
15. **Guarente, L., R. R. Yocum, and P. Gifford.** 1982. A *GAL10-CYC1* hybrid yeast promoter identifies the *GAL4* regulatory region as an upstream site. Proc. Natl. Acad. Sci. USA **79**:7410–7414.
16. **Hope, I. A., and K. Struhl.** 1985. *GCN4* protein, synthesized *in vitro*, binds *HIS3* regulatory sequences: implications for general control of amino acid biosynthetic genes in yeast. Cell **43**:177–188.
17. **Johnson, A. D., and I. Herskowitz.** 1985. A repressor (*MATalpha2* product) and its operator control expression of a set of cell type specific genes in yeast. Cell **42**:237–247.
18. **Khoury, G., and P. Gruss.** 1983. Enhancer elements. Cell **33**:313–314.
19. **Lawther, R. P., and T. G. Cooper.** 1975. Kinetics of induced and repressed enzyme synthesis in *Saccharomyces cerevisiae*. J. Bacteriol. **121**:1064–1073.
20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory,

Cold Spring Harbor, N.Y.

21. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74:**560–564.
22. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 403. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Olsen, J., S. Hahn, and L. Guarente.** 1987. Yeast *HAP2* and *HAP3* activators both bind to the *CYC1* upstream activation site, *UAS2* in an interdependent manner. Cell **51:**953–961.
24. **Rai, R., F. S. Genbauffe, and T. G. Cooper.** 1988. Transcriptional regulation of the *DAL5* gene in *Saccharomyces cerevisiae*. J. Bacteriol. **170:**266–271.
25. **Rai, R., F. S. Genbauffe, R. A. Sumrada, and T. G. Cooper.** 1989. Identification of sequences responsible for transcriptional activation of the allantoate permease gene in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **9:**602–608.
26. **Sarokin, L., and M. Carlson.** 1986. Short repeated elements in the upstream regulatory region of the *SUC2* gene of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **6:**2324–2333.
27. **Schule, R., M. Muller, C. Kaltschmidt, and R. Renkawitz.** 1988. Many transcription factors interact synergistically with steroid receptors. Science **242:**1418–1420.
28. **Smith, H. O., and M. L. Birstiel.** 1976. A simple method for DNA restriction site mapping. Nucleic Acids Res. **3:**2387–2398.
29. **Sumrada, R., and T. G. Cooper.** 1974. Oxaluric acid: a non-metabolizable inducer of the allantoin degradative enzymes in *Saccharomyces cerevisiae*. J. Bacteriol. **171:**1240–1247.
30. **Sumrada, R. A., and T. G. Cooper.** 1982. Isolation of the *CAR1* gene from *Saccharomyces cerevisiae* and analysis of its expression. Mol. Cell. Biol. **2:**1514–1523.
31. **Sumrada, R. A., and T. G. Cooper.** 1987. Ubiquitous upstream repression sequences control activation of the inducible arginase gene in yeast. Proc. Natl. Acad. Sci. USA **84:**3997–4001.
32. **Sumrada, R. A., and T. G. Cooper.** 1989. Nucleotide sequence of the *Saccharomyces cerevisiae* arginase gene (*CAR1*) and its transcription under various physiological conditions. J. Bacteriol. **160:**1078–1087.
33. **Sumrada, R., C. A. Zacharski, V. Turoscy, and T. G. Cooper.** 1978. Induction and inhibition of the allantoin permease in *Saccharomyces cerevisiae*. J. Bacteriol. **135:**498–510.
34. **Thiele, D. J., and D. H. Hamer.** 1986. Tandemly duplicated upstream control sequences mediate copper-induced transcription of the *Saccharomyces cerevisiae* copper-metallothionein gene. Mol. Cell. Biol. **6:**1158–1163.
35. **Turoscy, V., and T. G. Cooper.** 1982. Pleiotropic control of five eucaryotic genes by multiple regulatory elements. J. Bacteriol. **151:**1237–1246.
36. **Whitney, P. A., T. G. Cooper, and B. Magasanik.** 1973. Allophanate, the inducer of urea carboxylase and allophanate hydrolase in *Saccharomyces cerevisiae*. J. Biol. Chem. **248:**6203–6209.
37. **Yoo, H. S., F. S. Genbauffe, and T. G. Cooper.** 1985. Identification of the ureidoglycolate hydrolase gene in the *DAL* gene cluster of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **6:**3954–3964.
38. **Zinn, K., and T. Maniatis.** 1986. Detection of factors that interact with the human β-interferon regulatory region in vivo by DNAase I footprinting. Cell **45:**611–618.