Mutational Analysis of Upstream Activation Sequence 2 of the *CYC1* Gene of *Saccharomyces cerevisiae*: a *HAP2-HAP3*-Responsive Site

SUSAN L. FORSBURG AND LEONARD GUARENTE*

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

Received 31 August 1987/Accepted 3 November 1987

We analyzed upstream activation sequence 2 (UAS2), one of two independent UAS elements in the *CYC1* gene of *Saccharomyces cerevisiae*. Deletions and linker scanning mutations across the 87 base pairs previously defined as UAS2 showed two separate functional elements required for full activity. Region 1, from -230 to -200, contains the principal activation site and responds to the *trans*-acting regulatory loci *HAP2* and *HAP3*. A portion of region 1 is homologous to two other *HAP2-HAP3*-responsive UASs and includes the G→A transition mutation UP1, which increases UAS2 activity. This consensus sequence TNATTGGT bears striking similarity to several CAAT box sequences of higher cells. Region 2, from -192 to -178, substantially enhances the activity of region 1, yet has little activity by itself. These regions bind distinct proteins found in crudely fractionated yeast extracts.

Upstream activation sequences (UASs) in *Saccharomyces cerevisiae*, like mammalian enhancers, are required for regulation and efficient transcription of a variety of genes (for reviews, see references 5 and 6). Several UAS elements have been examined and shown to bind specific regulatory proteins, and the sequences necessary for this function have been identified. For example, the GAL4 protein binds a 17-base-pair (bp) dyad in the *GAL* UAS and activates transcription in the presence of galactose (3). A 9-bp dyad has been identified as the principal target sequence of the GCN4 protein (10, 12), which acts in the general control system regulating many amino acid biosynthetic genes (11).

In other cases, single UAS sites are more complex and appear to interact with multiple regulatory factors. For example, UAS1 of *CYC1* (encoding iso-1-cytochrome *c* [14]) and the UASs of *CYC7* (iso-2-cytochrome *c* [22, 25]) and *TRP1* (13) contain subsites which act synergistically to activate transcription. Such activation sites may allow a careful modulation of UAS activity or coordinate independent responses to different physiological signals. For example, the two principal subsites in the *CYC7* UAS mediate induction by heme and by carbon source, respectively (22, 26).

In addition to UAS1, transcription of *CYC1* is activated by a second, independent site, UAS2 (Fig. 1) (7). While basal transcription of *CYC1* under glucose conditions is activated by UAS1, induction in nonfermentable carbon sources occurs largely via UAS2 (7, 8). The two sites may be further distinguished by the separate regulatory loci that affect them. UAS1 is regulated by the HAP1 activator (7), which has been shown to bind the site *in vitro* (14, 18, 19). UAS2, in contrast, is activated by the products of the *HAP2* and *HAP3* loci (7, 20, 21). These regulatory genes are part of a global network modulating the expression of genes required for respiratory competence (7, 21) including *HEM1* (coding for δ -aminolevulinic synthase [T. Keng and L. Guarente, Proc. Natl. Acad. Sci. USA, in press]) and *COX4* (coding for subunit IV of cytochrome oxidase [J. C. Schneider and L. Guarente, unpublished data]). Cells with mutations of *HAP2*

or *HAP3* fail to activate UAS2 and are unable to grow on a nonfermentable carbon source (7, 21).

Previous analysis of UAS2 defined the boundaries of the site as lying within the interval -263 to -178 and identified a mutation at position -208 which increases UAS2 activity (UP1 [7]). The UP1 mutation changes the sequence TGGTTGGT (between -203 and -211) in wild-type UAS2 to TGATTGGT. *COX4* and *HEM1* bear the sequences TTATTGGT and TCATTGGT, respectively, in their UASs, giving a consensus of TNATTGGT. There is no further homology apparent between UAS2 and these other two UAS elements (Fig. 2).

With this framework in place, we carried out a more detailed analysis of UAS2. In this report, we describe the effect of deletions and linker scanning mutations on both native UAS2 and UAS2UP1 function *in vivo*. Surprisingly, we found that sequences critical for UAS2 activity, while including the TNATTGGT element described above, were much more extensive. The *HAP2-HAP3* effect maps to a subsite of these sequences.

MATERIALS AND METHODS

Strains and media. *S. cerevisiae* BWG1-7A (*MATa leu2-3 leu2-112 his4-519 ade1-100 ura3-52*) (8) and its derivatives *hap1-1* (7), *hap2-1* (7), and *hap3-1* (21) were used in this study. Cells were grown in synthetic or YEP medium supplemented with 2% glucose or lactate (24). Yeast transformations were done by the lithium acetate method (24).

β -Galactosidase assays. β -Galactosidase assays upon cells grown in synthetic glucose or lactate medium were done as described previously (17).

DNA isolation and manipulation. DNA manipulations were done by the method of Maniatis et al. (15), and enzymes were used according to the recommendations of the suppliers (New England BioLabs, Inc., Beverly, Mass., and Boehringer Mannheim Biochemicals, Indianapolis, Ind.). Fragments were purified from acrylamide or agarose by electroelution or with NA45 DEAE membrane as described by the supplier (Schleicher & Schuell, Inc., Keene, N.H.). *KpnI* linkers (sequence GGGTACCC) were purchased from New England BioLabs.

* Corresponding author.

FIG. 1. Organization of the *CYC1* upstream region. The *CYC1* leader contains two UASs, UAS1 centered at approximately -275 and UAS2 centered at approximately -200 . There are three functional TATA boxes (T) at -106 , -52 , and -22 and six distinct clusters of mRNA start sites (9). The upstream-most start site is termed $+1$. Numbering of the upstream region is according to Hahn et al. (9). Please note that this numbering is 3 bases different from that of the earlier papers (e.g., reference 7) owing to refinement in mapping the most upstream mRNA initiation site.

DNA sequencing. Fragments were cloned into M13mp19 and sequenced by the dideoxy chain termination method (23).

Primer extension. RNA isolation and primer extension were done by the method of Hahn et al. (9). The synthetic oligonucleotide primer used is homologous to a region of *lacZ* sequence in the *lacZ* fusion about 100 bases 3' to the most upstream *CYC1* start site.

Plasmid constructions. The starting $2\mu\text{m}$ vectors, pLGA-265 and pLGA-265UP1, have been described previously (7) (Fig. 3). Briefly, these are *CYC1-lacZ* fusions regulated by UAS2; there are no *CYC1* sequences upstream of a unique *SmaI* site at -263 , which is provided by the junction with the 3' end of the *URA3* gene. *CYC1* sequences extend from -263 to the fusion junction with *lacZ* and include UAS2 and the TATA region. These vectors contain a unique *XhoI* site at -178 (numbering from the upstream-most RNA start site) and a unique *SacI* site within the β -galactosidase gene. A UAS-less derivative with a *KpnI* linker (pSLF Δ 178K) was constructed by dropping out the 85-bp *SmaI-XhoI* fragment, filling in the ends with Klenow fragment, and ligating to a *KpnI* linker. This construction preserves all three restriction sites (*SmaI*, *KpnI*, *XhoI*).

Deletions. 5' deletions were constructed by cutting at *SmaI* of pLGA-265 and pLGA-265UP1 and treating with *Bal* 31 exonuclease (Fig. 3). The DNA was treated with Klenow fragment to blunt the ends and ligated to *KpnI* linkers. It was then cut with *KpnI* and *SacI*, and the fragment containing UAS2 was gel isolated and ligated to the backbone fragment from the *KpnI-SacI* digest of pSLF Δ 178K, so the deletions were unidirectional. The 3' deletions were treated similarly, the starting vectors in this case being cut with *XhoI*. We isolated the large *KpnI-SacI* fragment carrying the 3' deletions of UAS2 and ligated it to the smaller fragment from pSLF Δ 178K. Deletions were screened by restriction analysis. Those with endpoints mapping within UAS2 were sequenced and transformed into yeast cells.

Linker scans and insertions. Linker substitutions (16) were constructed by cutting the component deletions (with endpoints 7 to 9 bases apart) with *KpnI* and *SacI*, isolating the large backbone piece from the 3' deletions and the smaller fragment from the 5' deletions, and ligating the two together

(Fig. 3). If the deletion endpoints are closer together than 8 bases, a net insertion results; pSLFins1 was constructed this way and contains an insertion of the sequence ACGGG TACCC between -191 and -192 . These constructions were verified by restriction analysis and sequencing and were transformed into yeast cells. All deletion and linker substitution constructions maintain UAS2 on a convenient cassette flanked by unique *SmaI* and *XhoI* sites.

Gel electrophoresis DNA binding assays. Protein-DNA complexes were resolved on polyacrylamide gels as described previously (18). The 20- μl binding reaction mixture contained 1 mM dithiothreitol, 4 mM Tris hydrochloride (pH 8.0), 40 mM NaCl, 4 mM MgCl_2 , 5% glycerol, and 10 μg of proteins. This mixture was incubated at room temperature for 10 min and loaded onto a 4% polyacrylamide gel in TBE (90 mM Tris hydrochloride, 90 mM H_3BO_3 , 2.5 mM EDTA). The gels were run at 25 mA until the bromophenol blue had run to the base of the gel and were then dried and autoradiographed.

Extract preparation. Yeast cell extracts were prepared as described previously (18). Briefly, cells were grown to an A_{600} of 1.0, harvested by centrifugation, suspended in extraction buffer (200 mM Tris hydrochloride [pH 8.0], 400 mM $(\text{NH}_4)_2\text{SO}_4$, 10 mM MgCl_2 , 1 mM EDTA, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride, 7 mM 2-mercaptoethanol), and disrupted by agitation with glass beads. The extracts were centrifuged for 1 h at $10,000 \times g$; the supernatant was collected and precipitated with saturated $(\text{NH}_4)_2\text{SO}_4$ added to a final concentration of 50%. The protein was suspended in protein buffer (20 mM HEPES [*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid] [pH 8.0], 5 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 20% glycerol, and 7 mM 2-mercaptoethanol).

Probes for binding assays. UAS2 probes were prepared by digesting the constructs with *SmaI* and *XhoI* and end labeling at the *XhoI* site with Klenow fragment and $(\alpha\text{-}^{32}\text{P})\text{TTP}$ by standard methods (15). The probes were gel purified before use.

RESULTS

Analytical strategy. Deletions and substitution mutations of UAS2 sequences were constructed in plasmids with

FIG. 2. DNA sequence homology between UAS2UP1 and the UASs of *COX4* (J. C. Schneider and L. Guarente, unpublished data) and *HEM1* (Keng and Guarente, in press). The consensus sequence is boxed. *COX4* and *HEM1* are numbered from the first ATG in the coding sequence. The *COX4* sequence is inverted with respect to the coding sequence to allow the match.

CYC1-lacZ fusions activated by wild-type or UP1 forms of UAS2 (see Materials and Methods). 5' deletions beginning at -263 and 3' deletions beginning at -178 were constructed separately in both vectors. A *KpnI* linker was inserted at the deletion terminus. Selected deletions were recombined through the octamer *KpnI* linker to generate linker substitutions (16); most of these were within 1 bp of being exact 8-bp substitutions. 5' deletions in wild-type UAS2 were also constructed without the linker, to verify that the linker does not contribute to the phenotype of constructions bearing deletions (data not shown). Cells were grown under noninducing (i.e., glucose) or inducing (i.e., lactate) conditions; effects of UAS2 mutations on activity were determined for all mutants by assaying β -galactosidase levels and verified by analyzing RNA levels in selected constructions.

5' deletions of UAS2. A range of 5' deletions was constructed to allow definition of the 5' boundary of the site. The results were qualitatively the same in UAS2 and in UAS2UP1 (Table 1; Fig. 4). Deletion of sequences upstream of about -234 had no effect on β -galactosidase levels of wild type or UAS2UP1 in lactate. There was a sharp decline in the activity of both forms of the UAS as sequences between -234 and -210 were deleted. The remaining activity (1% of that of the intact UAS under lactate conditions) was abolished only when sequences between -192 and -185 were deleted. Thus, under inducing conditions, these deletions suggest that a critical portion of UAS2 has a 5' boundary at

FIG. 3. Structure of plasmids used in this study, constructed as described in Materials and Methods. All carry the yeast 2μ origin (on a 2.2-kilobase *EcoRI* fragment) and *URA3* marker, as well as the bacterial *Amp^r* gene and origin (*ori*) from pBR322. All plasmids contain the *CYC1-lacZ* fusion and *CYC1* TATA. pLGA-265 and pLGA-265UP1 contain, respectively, UAS2 and UAS2UP1 and are otherwise essentially identical. They each contain *CYC1* DNA up to -263, where the junction with the *URA3* gene forms a *SmaI* site. pSLF178K drops out the *SmaI-XhoI* fragment containing UAS2 and replaces it with a *KpnI* linker; this maintains the flanking *SmaI* and *XhoI* sites as well. 5' deletions have the *SmaI* site at the *URA3* junction abutting a *KpnI* linker and various amounts of 3' UAS2 sequence; similarly, the 3' deletions maintain the *XhoI* site abutting the *KpnI* linker, with various amounts of 5' UAS2 sequence. Linker scans retain both the *SmaI* and *XhoI* sites flanking UAS2, with the *KpnI* linker somewhere in between.

TABLE 1. Activity of 5' deletions in UAS2 and UAS2UP1^a

Deletion	β -Galactosidase activity ^b			
	BWG1-7A		<i>hap2</i>	<i>hap3</i>
	Glucose	Lactate	(glucose)	(glucose)
5' deletions in wild-type UAS2				
pLGA-265 ^c	5	300	1	1
pSLF178K	3	330	1	1
pSLF178K	3	280		
pSLF178K	2	250		
pSLF178K	2	70	1	1
pSLF178K	2	9	1	1
pSLF178K	2	5		
pSLF178K	2	4		
pSLF178K	2	5	2	2
pSLF178K	1	3	1	1
pSLF178K	1	1	1	1
pSLF178K	0.5	0.5	0.5	0.5
5' deletions in UAS2UP1				
pLGA-265UP1 ^c	100	600	2	2
pSLF178KUP1	90	590		
pSLF178KUP1	65	590		
pSLF178KUP1	65	600	1	1
pSLF178KUP1	50	580		
pSLF178KUP1	50	520		
pSLF178KUP1	45	540		
pSLF178KUP1	20	400		
pSLF178KUP1	2	150	0.5	0.5
pSLF178KUP1	2	80	1	1
pSLF178KUP1	2	15	1	0.5
pSLF178KUP1	2	5		
pSLF178KUP1	2	5		

^a Deletions with indicated endpoints were constructed as described in the text and transformed into the wild-type strain BWG1-7A or its *hap2* or *hap3* derivative. Only those constructions actually retaining the G-to-A transition at -208 are listed as UP1.

^b β -Galactosidase units are as described in reference 17.

^c Intact wild-type UAS2.

^d Deletes entire UAS2.

^e Intact UAS2UP1.

approximately -230 and raise the possibility that a second functional site exists around -190.

The wild-type UAS2 is scarcely active under glucose (uninduced) conditions, and the 5 U of activity from the starting plasmid pLGA-265 can be attributed in part to the UAS1 sequences contained between -263 and -244. When placed in a *hap1* strain (which specifically affects UAS1), pLGA-265 lost roughly half its glucose activity and was indistinguishable from Δ 244K (data not shown). The remaining activity may be attributed to the sequences in the vicinity of -190, as it was not affected until deletions crossed that point. In contrast, UAS2UP1 displayed a much higher basal activity in glucose than the wild-type site did (Table 1). The effect of 5' deletions in UAS2UP1 on this activity in glucose closely mimicked the effect of deletions in lactate (Fig. 4); the profile shows that glucose activity dropped steadily as deletions passed through the 5' region defined in lactate. The UP1 mutation thus appears to partially induce the site under glucose conditions. Conceivably, this mutation could increase the affinity of UAS2 for an activator that is limiting in glucose.

3' deletions in UAS2. 3' deletions corroborated the importance of the sequences in the -190 region (Table 2; Fig. 4). Deletion of sequences between -178 and -191 resulted in a loss of some 70% (UP1) to 90% (wild type) of activity in lactate. However, both wild-type and UP1 forms of these

FIG. 4. Deletion analysis of UAS2. Deletion endpoints are plotted against β -galactosidase activity, which is expressed as percent activity from the intact UAS2UP1 (for which 100% levels are 100 U in glucose and 600 U in lactate) or UAS2 (for which 100% levels are 5 U in glucose and 300 U in lactate). Symbols: $\blacktriangle$, $\bullet$, 5' deletions; $\triangle$, $\circ$, 3' deletions; $\blacktriangle$ and $\triangle$, activity under lactate conditions; $\bullet$ and $\circ$, activity under glucose conditions.

deletions were still regulated by carbon source; despite their diminished overall activity, they each were induced 10-fold in lactate medium. The upstream part of UAS2 thus contains the principal region required for carbon source response. A 3' deletion extending to -200 in UP1 did not result in any further loss of activity. The next deletion, to -211 , removed the consensus sequence (which lies between -210 and -203) and reduced levels of β -galactosidase even more. By -215 , even this modest remaining lactate activity was completely lost.

Under glucose conditions, the effects of the 3' UAS2UP1 deletions again clearly paralleled the effects seen in lactate. The very low level of activity of the wild-type UAS under glucose conditions and the high background again made a meaningful deletion analysis impossible.

Linker substitution analysis of UAS2. By constructing more precise linker scanning mutations in UAS2, we hoped to refine our analysis of sequences required for UAS2 activity and, further, to determine whether any sites in the -230 to -190 interval mediate negative control of UAS2 activity in glucose, perhaps by binding a repressor. The effects of linker substitution mutations on UAS2 expression are shown in Fig. 5 and Table 3. Representative linker scans were also integrated at the *URA3* locus and showed similar regulation in single copy to that found in high copy (data not shown). Several findings are apparent in this analysis. First, no substitution rendered the site constitutive in glucose, suggesting that negative control at UAS2 does not occur. Second, linker substitutions that exerted large effects on activity lay in three separate regions of UAS2UP1; substitutions that fell between these regions had small effects, if any, on expression of the site. The first critical sequence in

UAS2UP1 lies between -223 and -227 , at the 5' border of the region identified by deletion analysis as crucial for carbon source response. The second element includes the conserved sequence TNATTGGT found in the *HAP2-HAP3* coregulated genes. The third sequence lies near -190 , a region judged important by the deletion analysis described above. Substitutions of these critical sequences in both wild-type and UP1 backgrounds had drastic effects on activity.

The two substitutions ($\Delta 191-200$ UPK and $\Delta 217-223$ UPK) separating the three regions had less than a twofold effect on activity of UAS2UP1. In neither case did the sequence of the *KpnI* linker (GGGTACCC) fortuitously reconstitute all the UAS2 sequences substituted. Inspection of the sequence surrounding the linker scan $\Delta 217-223$ UPK revealed the presence of a short dyad, GCTC-N₇-GAGC. This linker substitution precisely substitutes the central seven bases, leaving intact the dyad itself. Deletion analysis suggests that the dyad is a functional unit. Specifically, deletion 3X211K, which removes the consensus sequence but retains the intact dyad, still induced fivefold, while deletion 3X213K, extending into the right half of the dyad, failed to induce (Table 2). We emerge with a refined picture of the carbon source response region of UAS2. It contains the consensus sequence and selected upstream sequences which include a small dyad.

A further mutation provides insight into the relationship between the carbon source-responsive portion of UAS2 and the -190 downstream component. An insertion of the sequence ACGGGTACCC between -192 and -191 separated the two regions by an additional 10 bp and had no effect on activity (pSLFins1, Table 3). We thus consider it unlikely

that a single regulatory protein would bind sequences that span both the consensus element and the -190 region. Rather, we suggest that separate proteins bind to the carbon source and downstream portions of UAS2, a matter investigated below.

RNA analysis. To verify that the effects of UAS2 mutations on β -galactosidase accurately reflect effects on transcription, we carried out primer extension analysis on the linker substitutions derived from UAS2UP1. RNA levels were determined for three representative scans, showing low ($\Delta 203$ -211K), moderate ($\Delta 220$ -227UPK), and high ($\Delta 229$ -237UPK) levels of β -galactosidase activity (Fig. 6). It is apparent that RNA levels are proportional to activity as determined by β -galactosidase assays.

Effect of mutations in *HAP2* and *HAP3* on deletions and substitutions. To map that part of the UAS through which *HAP2* and *HAP3* mediate their effects, we transformed a range of deletions (5' and 3' in both wild-type and UP1 backgrounds) and linker substitutions into *hap2-1* and *hap3-1* mutants (Tables 1, 2, and 3). Similar results were found upon transformation into strains with disruptions of *HAP2* and *HAP3* (data not shown). UAS2UP1 retained only low levels of activity in the *hap2* and *hap3* mutants, approximately 2 units compared with 100 units in wild-type cells. The only linker substitutions with a similar low level of expression in a *Hap*⁺ strain were those affecting the conserved sequence, suggesting that this sequence is at least part of the *HAP2*-*HAP3*-responsive element. The activity of constructions that perturb the consensus was not additionally diminished in a *hap2* or *hap3* strain. However, expression of all other constructions declined further in the mutant strains. For the 5' and 3' deletion plasmids, activity in a *hap2* or *hap3* strain declined until the consensus was deleted, providing further evidence that the consensus element is indeed part of the *HAP2*-*HAP3*-responsive site. Also, expression of the 5' deletions which retain only the down-

TABLE 2. Activity of 3' deletions in UAS2 and UAS2UP1^a

Deletion	β -Galactosidase activity ^b			
	BWG 1-7A		<i>hap2</i>	<i>hap3</i>
	Glucose	Lactate	(glucose)	(glucose)
3' deletions in wild-type UAS2				
pLGA-265 ^c	5	300	1	1
pSLF3X189K	3	30		
pSLF3X211K	2	10	2	2
pSLF3X213K	2	2		
pSLF3X215K	1	1	1	1
pSLF3X219K	1	1		
pSLF3X223K	1	1		
pSLF3X227K	1	1		
pSLF3X237K	1	1		
pSLF3X242K	1	1	1	1
pSLF3X252K	1	1		
3' deletions in UAS2UP1				
pLGA-265UP1 ^d	100	600	2	2
pSLF3X192UPK	12	150		
pSLF3X200UPK	10	150	0.5	0.5

^a Deletions with indicated endpoints were constructed as described in the text and transformed into the wild-type strain BWG1-7A or its *hap2* or *hap3* derivative. Only those constructions actually retaining the G-to-A transition at -208 are listed as UP1.

^b β -Galactosidase units are as described in reference 17.

^c Intact wild-type UAS2.

^d Intact UAS2UP1.

stream region was not *HAP2*-*HAP3* dependent. Because the *hap2* and *hap3* strains are petite and may not be grown in lactate, it was not possible to evaluate the effect of these mutations on wild-type UAS2.

DNA binding assays. The mutational analysis presented above raised the possibility that different proteins might bind to different parts of UAS2. If such binding was noncooper-

FIG. 5. Linker substitution analysis of UAS2. Linker scans are represented by a horizontal bar (■) indicating the position of the exact substitution with the linker sequence GGGTACCC. β -Galactosidase activity of each construction is plotted for glucose (□) or lactate (■) conditions as the percent activity of the intact UAS2UP1 (for which 100% levels are 100 U in glucose and 600 U in lactate) or wild-type UAS2 (for which 100% levels are 5 U in glucose and 300 U in lactate).

TABLE 3. Activity of linker substitutions in UAS2 and UAS2UP1^a

Linker substitutions	β-Galactosidase activity ^b			
	BWG 1-7A		hap2 (glucose)	hap3 (glucose)
	Glucose	Lactate		
Substitutions in wild type				
pLGA-265 ^c	5	300	1	1
pSLFΔ203-211	3	12	2	2
pSLFΔ209-216	3	12	1	0.5
pSLFΔ220-227	2	15	2	2
pSLFΔ234-242	3	300	2	2
pSLFΔ244-252	4	320		
Substitutions in UAS2UP1				
pLGA-265UP1 ^d	100	600	2	2
pSLFΔ185-191UPK	15	120	0.5	0.5
pSLFΔ192-200UPK	60	500	5	3
pSLFΔ209-216UPK	3	12	1	1
pSLFΔ217-223UPK	30	500	1	1
pSLFΔ220-227UPK	5	90	2	2
pSLFΔ229-237UPK	70	600	2	2
pSLFΔ235-242UPK	70	590		
pSLFΔ244-252UPK	90	600		
Insertion (pSLFins1) ^e	5	300		

^a Linker substitutions with indicated endpoints were constructed as described in the text and transformed into the wild-type strain BWG1-7A or its *hap2* or *hap3* derivative. Only those constructions actually retaining the G-to-A transition at -208 are listed as UP1.

^b β-Galactosidase units are as described in reference 17.

^c Intact wild-type UAS2.

^d Intact UAS2UP1.

^e pSLFins1 is described in the text.

ative, these proteins could give rise to distinct complexes. To visualize such complexes, we employed a gel retardation binding assay (Fig. 7) using wild-type and UP1 forms of UAS2, as well as two deletion mutants. The probes used contained UAS2 sequences between -263 and -178 (UAS2 and UAS2UP1), -263 and -191 (3X191UPK), or -220 and -178 (Δ220UPK). Each DNA probe was bound to a crude extract from glucose- and lactate-grown cells. Three distinct complexes appeared, which we call A, B, and C. A and B were not affected by carbon source nor by *hap2* or *hap3* (J. Olesen, S. Hahn, and L. Guarente, Cell, in press). The appearance of C required the UP1 mutation in UAS2 and extracts prepared from lactate-grown cells. Two observations indicate that band C requires both the consensus sequence and the sequences upstream. First, formation of band C depended on the UP1 mutation in the consensus element, and second, band C was not seen in a construction that removed part of the upstream sequences (and part of the small dyad) without affecting the consensus (Δ220UPK). The slightly altered mobility of complexes A and B found with this construct may be due to the smaller size of this probe relative to the others.

Band A, on the other hand, depended only on the downstream sequence element, since the 3' deletion removing that element prevented its formation (Fig. 6, lanes 5 and 6). Since band B is found in all cases, we cannot determine its sequence requirements from this experiment. We conclude from these in vitro experiments that there are at least two specific protein complexes interacting with UAS2. These complexes appear to bind independently of one another. The binding of one complex (C), seen only in lactate, requires both the consensus element and sequences upstream, while

FIG. 6. Primer extension of representative linker substitution constructions. The primer hybridized to the *lacI* portion of the *lacZ* fusion. RNA was extracted from lactate-grown BWG1-7A cells containing the indicated plasmids, with β-galactosidase activity as follows: UAS2UP1 (lanes 1 and 2), 580 U; Δ203-211K (lanes 3 and 4), 10 U; Δ220-227UPK (lanes 5 and 6), 95 U; Δ229-237UPK (lanes 7 and 8), 520 U. Samples of 10 μg (lanes 1, 3, 5, and 7) and 40 μg (lanes 2, 4, 6, and 8) of total RNA were assayed for each construction. The reactions were sized against pBR322 cut with *MspI* (lane M). The typical six clusters of *CYC1* start sites are numbered at the right according to Hahn et al. (9).

FIG. 7. Detection of protein-DNA complexes by gel retardation assays with crude yeast extracts. Extracts from wild-type cells grown in glucose (G lanes: 1, 3, 5, and 7) or lactate (L lanes: 2, 4, 6, and 8) were assayed for binding to radiolabeled UAS2 fragments. Binding reactions were done in the presence of 1 μg of salmon sperm DNA. Lanes 1 and 2 contain wild-type UAS2 (from pLGA-265). Lanes 3 and 4 contain UAS2UP1 (from pLGA-265UP1). Lanes 5 and 6 contain the 3' deletion pSLF3X191UPK, which deletes the downstream region (region 2). Lanes 7 and 8 contain the 5' deletion pSLFΔ220UPK, which deletes half the upstream region (region 1). A, B, and C are defined in the text.

the second complex (A) requires the downstream region for binding.

DISCUSSION

In this report, we present a detailed mutational analysis of the upstream activation site UAS2 of the *S. cerevisiae* *CYC1* gene. The activity of this site in vivo requires the products of the *HAP2* and *HAP3* regulatory genes and is induced by nonfermentable carbon sources. Our analysis indicates that UAS2 contains two separate regions. Region 1 mediates the carbon source induction, and region 2 augments this activity. The UP1 mutation, a G→A transition at -208 which increases UAS2 activity in vivo, lies within region 1. This mutation also increases the UAS2 homology to the consensus TNATTGGT found in other genes activated by *HAP2-HAP3* such as *COX4* (J. C. Schneider and L. Guarente, unpublished data) and *HEM1* (Keng and Guarente, in press) (Fig. 2). We note a striking homology between this consensus and the CAAT box of promoters from genes of higher cells (1, 2, 4). For example, the sea urchin H2B1 gene contains the sequence GCGTTGATTGGT (with the CAAT box on the opposite strand [1]) which is perfectly homologous to bases -203 to -214 of UAS2UP1.

Because UAS2UP1 is fully dependent on *HAP2* and *HAP3* for activity and is affected by deletions and linker scanning mutations (see below) much as wild-type UAS2 is, we believe that the UP1 mutation acts to increase the affinity of UAS2 for its activator(s). Further, since the effect of the UP1 mutation is much greater in glucose (20-fold increase over wild-type levels) than in lactate (2-fold increase), it is likely that binding of the activator is regulated by carbon source and is severely limiting in glucose.

5' and 3' deletion analysis, combined with linker scanning mutagenesis, allowed delineation of the two regions critical for UAS2 activity (Fig. 8). Region 1 is located from -234 to -203 and is required for induction of UAS2 activity in a nonfermentable carbon source such as lactate. Mutations affecting the region 1 consensus element have very strong down phenotypes. However, sequences important in the activity of region 1 extend about 20 bp upstream of the consensus element. Deletions and linker scans affecting the sequence upstream of the consensus show greatly reduced activity. A short dyad lies within these upstream sequences; a substitution of the sequences between the two halves of the dyad has minimal phenotype, while deletions or scans that impinge on either half affect activity more strongly. An additional finding suggesting a role for the dyad is that a 3' deletion destroying the consensus and all downstream sequences but leaving the dyad intact still has weak activity and, significantly, still is induced in lactate.

By introducing various UAS2 mutants into *hap2* or *hap3* mutants, we determined that the *HAP2-HAP3* activation system exerts its effects through region 1. The consensus element is thus at least a part of the *HAP2-HAP3*-responsive site.

The downstream portion of the UAS (region 2, between -192 and -178) has minimal activity by itself (about 1% of that of the wild-type site in lactate) but acts synergistically to modulate the activity of the upstream region. A linker substitution or 3' deletion in region 2 reduces UAS2 activity in lactate by 70 to 90%. The modest activity from region 2 in isolation is not affected by mutating *HAP2* or *HAP3*.

Further confirmation of the distinct functionality of the two components of UAS2 comes from the in vitro analysis by gel retardation DNA binding assays. Two specific, inde-

FIG. 8. Model for UAS2. There are two components. Region 1 consists of the principal carbon source response site; it contains the UP1 mutation and the only homology to the *HAP2-HAP3*-coregulated genes *COX4* and *HEM1*, as well as an additional 20 bases of DNA upstream of this consensus. *HAP2* and *HAP3* mediate their effects via region 1. Region 2 is required for maximum activity from region 1, but has little function by itself. Each region binds in vitro a distinct protein complex.

pendent complexes bind to UAS2. One complex requires region 2 sequences and has been shown in work to be presented elsewhere to be *HAP2* and *HAP3* independent (Olesen et al., in press). We have no proof at this point that the region 2-specific complex seen in vitro has in vivo significance. The second complex observed requires region 1 sequences, including both the consensus element and upstream dyad, and is only seen in extracts from lactate-grown cells. In addition, the formation of this complex in our assay requires the UP1 form of UAS2. In experiments to be described elsewhere, we show that this complex contains both *HAP2* and *HAP3* (Olesen et al., in press).

Since each UAS2 complex may be formed in the absence of the other, we conclude that binding of factors to region 1 and region 2 is not cooperative. This conclusion is intriguing, since it suggests that the synergy of the site is not determined by cooperative binding, but by cooperativity in the activation process. A similar situation was found in the case of *CYC7* activation, which is mediated by *HAP1* and a factor, augmenting activity, that binds downstream (22).

It is curious that other genes activated by *HAP2-HAP3*, while bearing homology to the consensus element, show little homology to the upstream dyad sequence (Fig. 2). This finding could mean that most of the important sequences within region 1 are within the consensus element, while only scattered bases within the upstream sequence are utilized in binding. This view would postulate that the dyad sequence occurs fortuitously. Alternatively, the sequence arrangements of UAS2 and other *HAP2-HAP3*-regulated genes may predict recognition by a family of protein complexes with a common core. By this model, the core would specify recognition of the consensus element while additional, UAS-specific factors would recognize the upstream sequences.

In summary, UAS2, like UAS1 and the *CYC7* UAS, consists of at least two functional components which interact synergistically to provide full expression. One distinct site mediates carbon source response, while the second site augments the activity of the first. Distinct protein complexes appear to bind to each region. This combinational design may allow more precise modulation of the expression of UAS2 over a range of conditions and is likely to be a common means of organizing regulatory sequences.

ACKNOWLEDGMENTS

We thank J. Olesen, S. Hahn, and K. Pfeifer for helpful advice on the gel binding assay and J. Olesen for crude extracts. We also thank K. Pfeifer and M. Haldi for critical reading of the manuscript.

This work was supported by Public Health Service grant 5 RO1 GM30454-05 to L.G. from the National Institutes of Health. S.L.F.

was supported by a National Science Foundation predoctoral fellowship.

LITERATURE CITED

1. Barberis, A., G. Superti-Furga, and M. Busslinger. 1987. Mutually exclusive interaction of the CCAAT-binding factor and of a displacement protein with overlapping sequences of a histone gene promoter. *Cell* 50:347-359.
2. Breathnach, R., and P. Chambon. 1981. Organization and expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 50:349-353.
3. Giniger, E., S. M. Varnum, and M. Ptashne. 1985. Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell* 40:767-774.
4. Gluzman, Y. 1985. Eukaryotic transcription: the role of *cis*- and *trans*-acting elements in initiation. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. Graves, B. J., P. F. Johnson, and S. L. McKnight. 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. *Cell* 44:565-576.
6. Guarente, L. 1984. Yeast promoters: positive and negative elements. *Cell* 36:799-800.
7. Guarente, L., B. Lalonde, P. Gifford, and E. Alani. 1984. Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *S. cerevisiae*. *Cell* 36:503-511.
8. Guarente, L., and T. Mason. 1983. Heme regulates transcription of the *CYC1* gene of *S. cerevisiae* via an upstream activation site. *Cell* 32:1279-1286.
9. Hahn, S., E. Hoar, and L. Guarente. 1985. Each of three "TATA elements" specifies a subset of the transcription initiation sites at the *CYC1* promoter of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 82:8562-8566.
10. Hill, D. E., I. A. Hope, J. P. Macke, and K. Struhl. 1986. Saturation mutagenesis of the yeast *his3* regulatory site: requirements for transcriptional induction and for binding by the GCN4 activator protein. *Science* 234:451-457.
11. 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.
12. Hope, I., and K. Struhl. 1985. GCN4 protein, synthesized *in vitro*, binds HIS3 regulatory sequences: implications for the general control of amino acid biosynthetic genes in yeast. *Cell* 43:177-188.
13. Kim, S., J. Mellor, A. J. Kingsman, and S. M. Kingsman. 1986. Multiple control elements in the *TRP1* promoter of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 6:4251-4258.
14. Lalonde, B., B. Arcangioli, and L. Guarente. 1986. A single *Saccharomyces cerevisiae* upstream activation site (UAS1) has two distinct regions essential for its activity. *Mol. Cell. Biol.* 6:4690-4696.
15. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. McKnight, S. L., and R. Kingsbury. 1982. Transcriptional control signals of a eukaryotic protein coding gene. *Science* 217:316-324.
17. Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. 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.
19. Pfeifer, K., T. Prezant, and L. Guarente. 1987. Yeast HAP1 activator binds to two upstream sites of different sequences. *Cell* 49:19-27.
20. Pinkham, J. L., 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.
21. Pinkham, J. L., J. T. Olesen, and L. Guarente. 1987. Sequence and nuclear localization of the *Saccharomyces cerevisiae* HAP2 protein, a transcriptional activator. *Mol. Cell. Biol.* 7:578-587.
22. Prezant, T., K. Pfeifer, and L. Guarente. 1987. Organization of regulatory region of the yeast *CYC7* gene: multiple factors are involved in regulation. *Mol. Cell. Biol.* 7:3252-3259.
23. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
24. Sherman, F., G. R. Fink, and J. B. Hicks. 1983. Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. Wright, C. F., and R. S. Zitomer. 1984. A positive regulatory site and a negative regulatory site control the expression of the *Saccharomyces cerevisiae* *CYC7* gene. *Mol. Cell. Biol.* 4:2023-2030.
26. Zitomer, R. S., J. W. Sellers, D. W. McCarter, G. A. Hastings, P. Wick, and C. V. Lowry. 1987. Elements involved in oxygen regulation of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 7:2212-2220.