Curr Genet (2001) 39: 156–165
DOI 10.1007/s002940100201

ORIGINAL PAPER

George K. van der Merwe · Hennie J. J. van Vuuren
Terrance G. Cooper

# *Cis*-acting sites contributing to expression of divergently transcribed *DAL1* and *DAL4* genes in *S. cerevisiae*: a word of caution when correlating *cis*-acting sequences with genome-wide expression analyses

Received: 6 November 2000 / Accepted: 8 February 2001 / Published online: 10 April 2001
© Springer-Verlag 2001

**Abstract** Correlating genome-wide expression profiles with sequence searches of promoter regions is being used as a technique to identify putative binding sites for *trans*-acting factors or to refine consensus sequences of those already known. To evaluate the limitations of such an approach in our studies of GATA-mediated transcription in *Saccharomyces cerevisiae*, we identified the relative contributions made to *DAL1* and *DAL4* expression by each of five Gln3p-, and/or Gat1p-, and three Dal82p-binding site homologous sequences situated in the 829-bp intergenic region separating these highly related, divergently transcribed genes. Our data suggest that although the correlation of repeated sequences or sequence homologies appearing within promoter regions with expression profiles obtained from genome-wide transcription analyses can provide useful starting points for analyses of *cis*-acting sites, significant limitations and possibilities for misinterpretation also abound.

## Introduction

Cluster analysis of genome-wide gene expression data is currently being used both to identify genes controlled by known transcription factors and to identify and generate consensus sequences for transcription factor-binding sites (Kruglyak and Tang 2000; Lyons et al. 2000). Such analyses usually designate 1 kb upstream of the clustered gene's ATG as the search target, following the rationale that "two genes that are controlled by a single regulatory system should have similar expression patterns in any data set" (Kruglyak and Tang 2000). Although such analyses can yield useful information, they are also potentially subject to significant problems and limitations, particularly in cases of divergently transcribed genes sharing small intragenic regions. The principal difficulties are knowing whether sequences, observed by homology or repeated appearance in an upstream region, are: (1) functioning in vivo, (2) shared equally by the divergently transcribed genes, and (3) influenced by the operation of other *trans*-acting factors.

A brief survey of the literature yielded analyses of 14 divergently transcribed genes (Angermayr and Bandlow 1997; Bell et al. 1995, 1997; Friesen et al. 1997; Hahn et al. 1988; Halfter et al. 1989; Johnston and Davis 1984; Kraakman et al. 1989; Kruglyak and Tang 2000; Liu and Xiao 1997; Osley et al. 1986; Schlapp and Rodel 1990; Siliciano and Tatchell 1984; Struhl 1985; Thuriaux et al. 1995); and the genome contains many more that are unstudied. In most cases, the analysis identified relatively large DNA fragments (greater than 40–100 bp) that support regulated gene expression in heterologous vectors. For a few (e.g., *GAL1–GAL10*, *PET56–HIS3*, *MAL6T–MAL6S*), the analyses have been more comprehensive, although not exhaustive (Bell et al. 1995; Johnston and Davis 1984; Struhl 1985). For example, *GAL1* and *GAL10* are reported to be regulated in common by four Gal4p-binding sites situated in a 75-bp region between them (West et al. 1984).

To evaluate problems potentially associated with making correlations such as those mentioned above, we analyzed, as a model, the expression of two closely related, divergently transcribed genes, *DAL1* and *DAL4* (encoding allantoinase and allantoin permease, respectively; Cooper 1996). They share an intergenic region (829 bp) roughly the same size as *GAL1–GAL10* (680 bp) and eight sequences qualifying as homologous to known allantoin pathway transcription factor-binding sites.

Communicated by E. Dubois

T. G. Cooper (✉)
Department of Microbiology and Immunology,
University of Tennessee, Memphis, TN 38163, USA
E-mail: tcooper@utmem.edu

G. K. van der Merwe · H. J. J. van Vuuren
BC Wine Research Center, Faculty of Agricultural Sciences,
University of British Columbia, Vancouver V6T 1Z4,
British Columbia, Canada

Two types of *cis*-acting elements are responsible for regulated *DAL* gene expression (see Cooper 1996; Hofman-Bang 1999; ter Schure et al. 2000; Wiame et al. 1985 for comprehensive reviews of the GATA-transcription factor literature): (1) upstream activating sequence (UAS) $UAS_{NTR}$ elements and (2) upstream induction sequence (UIS) $UIS_{ALL}$ elements. $UAS_{NTR}$ elements are dodecanucleotides with the sequence GATAA at their core (Bysani et al. 1991) that are binding sites both for the transcriptional activators Gln3p and/or Gat1p/Nil1p (Blinder and Magasanik 1995; Cunningham et al. 1996) and for the competing GATAA-binding repressor protein Dal80p (Cunningham and Cooper 1993; Cunningham et al. 1994). $UIS_{ALL}$ elements are dodecanucleotides that are binding sites for Dal82p which, along with other protein(s), is responsible for allophanate-induced gene expression (Dorrington and Cooper 1993; Van Vuuren et al. 1991). The ability of two of the $UIS_{ALL}$ elements situated between *DAL1* and *DAL4* to bind Dal82p has been measured (Dorrington and Cooper 1993). The inducer of *DAL* gene expression is allophanate, the last intermediate in the pathway or its non-metabolized analogue oxalurate (OXLU; Cunningham and Cooper 1993). There are five $UAS_{NTR}$-homologous and three $UIS_{ALL}$-homologous sequences in the *DAL1–DAL4* intergenic region that would qualify for inclusion in genome-wide correlations such as those mentioned above.

In this work, we determined the relative contributions of each $UAS_{NTR}$- and $UIS_{ALL}$-homologous sequence in the *DAL1–DAL4* intragenic region. The information obtained argues that even when analyzing highly related, divergently transcribed genes, whose transcription is supported by well characterized *cis*-acting elements, it is difficult to draw rigorous conclusions about: (1) which of the *cis*-acting element homologous sequence(s) found in a promoter region are actually responsible for the observed transcription, (2) their relative contributions to the overall transcription profile, and (3) the significance that should be attached to their presence in generating/refining a consensus sequence or identifying candidate sequences mediating an observed form of regulation in the absence of detailed biochemical analysis.

## Materials and methods

### Strains and media

All *Saccharomyces cerevisiae* strains used in this research are isogenic derivatives of the wild-type TCY1: TCY1 (*Matα ura3 lys2*), TCY17 (*Matα ura3 lys2 dal80Δ::hisG*), RR91 (*Matα ura3 lys2 gln3Δ::hisG*), HEY6 (Matα *ura3 lys2 dal81*Δ::*hisG*), and SS400 (*Matα ura3 lys2 trp1 dal82Δ::TRP1*). We also used *Escherichia coli* strain *DH5α* F′/endA1 hsdR17($r_K^- m_K^+$) supE44 thi1 recA1 gyrA (*Nal*$^r$) relA1 Δ(lacZYA-argF)$_{U169}$ (m80lacZΔM15). Yeast cultures for β-galactosidase and Northern blot analyses were grown in yeast nitrogen base (YNB) medium (0.17% YNB without amino acids or ammonium sulfate; Difco Laboratories, Detroit, Mich.), supplemented with 2% glucose, amino acids required to complement auxotrophies, and either 0.1% glutamine (repressed) or proline (derepressed) as sole nitrogen source. Gratuitous inducer, OXLU, was added to proline-containing medium (final concentration, 0.5 mM).

### Plasmid construction and PCR

A *CEN*-based *lacZ* reporter plasmid, for analyzing the expression supported by wild-type and mutant alleles of the *DAL1–DAL4* intergenic promoter region, was constructed by cloning a *Bam*HI linker (5′-CGCGGATCCGCG-3′) into the *Sma*I site of pHP41 (Park et al. 1992) to yield pVAN1. *Bam*HI digestion of pVAN1 yielded a 10-kb fragment which was isolated and re-circularized to yield pVAN2, which served as the parent plasmid in all β-galactosidase assays.

PCR-based methods were used to create deletion and substitution mutations in the *DAL1–DAL4* intergenic promoter region. Primers GK2 (5′-GCGCGGATCCGGAAG TGATGGCATT-GATAGGCATC-3′), and GK3 (5′-GCGCGGATCCAGCACTT-AGA GCGTCGTTAGCCATT-3′) were used to synthesize a fragment covering the nucleotides from +24 *DAL1* to *DAL4* +24, thereby allowing in-frame fusion of either the *DAL1* ATG or the *DAL4* ATG to the *lacZ* gene of pVAN2. The other primers we used are listed in Table 1.

The strategy used to mutate a specific potential *cis*-acting element, the template and primer combinations used, and the plasmids created are presented in Table 2. Site-directed mutations were constructed according to Viljoen et al. (1999). Heat-stable DNA polymerase PWO (Roche Molecular Biochemicals) was used in all PCR reactions. Reaction conditions and amplification programs were as prescribed by the manufacturer. All PCR products were digested with *Bam*HI and cloned into pVAN2. The integrity of all DNA fragments synthesized by PCR and in-frame fusions was confirmed by sequence analysis. The mutations introduced into each mutated promoter construct are listed in Table 2.

### Yeast and bacterial transformation

Yeast (Geitz et al. 1992) and bacterial (Inoue et al. 1990) transformation procedures have been described previously.

### β-Galactosidase assays

β-Galactosidase assays were performed essentially as described (Smart et al. 1996), except that we analyzed 10 ml of culture instead of 25 ml. Assays were performed in duplicate and from at least two independent yeast transformations. Data from duplicate assays generally varied less than 5% and from repeated transformations less than 20%. Enzyme activities are expressed in Miller units (Miller 1972), but are based on 10 ml of culture.

### Northern blot analysis

Total RNA was isolated from cultures grown to mid-log phase ($A_{600} = 1.0$; Ausubel et al. 1994). Poly(A)$^+$ RNA was isolated using the PolyATtract mRNA isolation system III (Promega), according to the manufacturer's recommendations. Samples of Poly(A)$^+$ RNA were resolved on 1.2% agarose-formaldehyde gels and transferred to Genescreen Plus 66 nylon membranes (NEN Research Products, Dupont). Double-stranded DNA probes used in the Northern blot analyses were synthesized by PCR using the oligonucleotides DAL1-5 (5′-CTGGCATCAATGAAAGC-3′) and DAL1-3 (5′-CTGCAGCAATACACAAA-3′) for *DAL1*, DAL4-5 (5′-ATGGCTAACGACGCTCT-3′) and DAL4-3 (5′-TATGAC-ACAATAGATGT-3′) for *DAL4*, and H4-5 (5′-GGCCGGATC-CATGTCCGGTAGAGGTAAAGG-3′) and H4-3 (5′-GGCCGA-ATTCTTAACCACCGAAACCGTATAAGG-3′) for *H4*. DNA

**Table 1** Oligonucleotides used in this research. Applications are relative to *DAL1* ATG (+1), unless otherwise indicated. *Italics* indicate mutations introduced. *Bold* indicates restriction sites used for cloning and introduction of mutations

| Primer | Sequence | Application |
|---|---|---|
| GK2 | 5′-GCGC**GGATCC**GGAAGTGATGGCATTGATAGGCATC-3′ | *DAL1*-ATG *lacZ* fusion |
| GK3 | 5′-GCGC**GGATCC**AGCACTTAGAGCGTCGTTAGCCATT-3′ | *DAL4*-ATG *lacZ* fusion[a] |
| Deletion analysis | | |
| GK7 | 5′-GCGC**GGATCC**GGGACAATAGAATCGAAACATGC-3′ | –543 of *DAL4* ATG[a] |
| GK8 | 5′-GCGC**GGATCC**AGCGGTCAATCCATCCTATTA-3′ | –560 of *DAL1* ATG |
| GK20 | 5′-GACT**GGATCCC**TGCATGTTTCGATTCTATT-3′ | –307 of *DAL1* ATG |
| $UAS_{NTR}$ site-directed mutation analysis | | |
| GK25 | 5′-TTGCGGTGCTT**AGACGT**CTATATAGAGGAG-3′ | Mutate –188 to –193 |
| GK26 | 5′-CTCCTCTATATA**GACGTC**TAAGCACCGCAA-3′ | Mutate –188 to –193 |
| GK11 | 5′-ACC**AAGCTT**AGATACCCTCGAGCTGCATGT-3′ | Mutate –357 to –362 |
| GK12 | 5′-TCT**AAGCTT**GGTATTACTTTCTTATCAATG-3′ | Mutate –357 to –362 |
| GK13 | 5′-TT**CCATGG**AAAGTAATACCGATAAGAGATA-3′ | Mutate –384 to –389 |
| GK14 | 5′-TT**CCATGG**AATGAAAAATTTCTGCCAGGGA-3′ | Mutate –384 to –389 |
| GK15 | 5′-TC**CAATTG**GCAACTAGATTAGAGGCGCTAT-3′ | Mutate –474 to –479 |
| GK16 | 5′-GC**CAATTG**GAATGTGTATGTGTAATTGAAG-3′ | Mutate –474 to –479 |
| GK17 | 5′-AC**GAATTC**GACGTGACAGCAAAGCGGTCAA-3′ | Mutate –574 to –579 |
| GK18 | 5′-TC**GAATTC**GTTCCTTTAAAGATTGTGTCCA-3′ | Mutate –574 to –579 |
| $UIS_{ALL}$ site-directed mutation analysis | | |
| GK33 | 5′-GATC*ACTAG***CAATTG**GCTTAATTATCTATATAGAGG-3′ | Mutate –197 to –210 |
| GK34 | 5′-GATC**CAATTG***CTAGT*GAACCACTTCTCCTGATTAAG-3′ | Mutate –197 to –210 |
| GK37 | 5′-GATC**GAATTC***ATATG*CCCTGGCAGAAATTTTTCATT-3′ | Mutate –399 to –411 |
| GK38 | 5′-GATC*CATAT***GAATTC**GCTTTTTTTCCGGCCATCCTTA-3′ | Mutate –399 to –411 |
| GK41 | 5′-GATC*TGATCA***ATGCAT**GTATGCGACAGCGAGTAAG-3′ | Mutate –448 to –460 |
| GK42 | 5′-GATC**ATGCAT***TGATCA*CTAATCTAGTTGCGATAAGG-3′ | Mutate –448 to –460 |

[a] Relative to *DAL4* ATG (+1)

**Table 2** PCR strategies to construct various deletion and substitution mutations in the *DAL1–DAL4* intergenic region. Sequence coordinates are relative to the *DAL1* ATG

| Sequences analyzed | Template and primers for PCR | Mutation | Plasmids created DAL1 and DAL4 |
|---|---|---|---|
| Wild-type promoter | pTC12[a]; GK2 and GK3 | None | pGV1 and pGV2 |
| $UAS_{NTR}$-directed mutations | | | |
| *GATA1* (–188 to –193) | pTC12; GK2/GK25 and GK26/GK3 | ATTATC → gacgTC | pGV3 and pGV4 |
| *GATA2* (–357 to –362) | pTC12; GK2/GK11 and GK12/GK3 | GATAAG → aAgctt | pGV5 and pGV6 |
| *GATA3* (–384 to –389) | pTC12; GK2/GK13 and GK14/GK3 | GATAAG → ccatgG | pGV7 and pGV8 |
| *GATA4* (–474 to –479) | pTC12; GK2/GK15 and GK16/GK3 | CTTATC → CaatTg | pGV9 and pGV10 |
| *GATA5* (–574 to –579) | pTC12; GK2/GK17 and GK18/GK3 | CTTATC → gaatTC | pGV11 and pGV12 |
| | pGV8[b]; GK2/GK11 and GK12/GK3 | Combine *gata2* and *3* mutations | pGV13 and pGV14 |
| | pGV12[b]; GK2/GK15 and GK16/GK3 | Combine *gata4* and *5* mutations | pGV15 and pGV16 |
| $UIS_{ALL}$-directed mutations | | | |
| *UIS6* (–197 to –210) | pTC12; GK2/GK33 and GK34/GK3 | CAAAATTGCGGTGC → CActAgcaattgGC | pGV21 and pGV22 |
| *UIS7* (–399 to –411) | pTC12; GK2/GK37 and GK38/GK3 | GGGCGCATTTTCC → GaattcaTaTgCC | pGV27 and pGV28 |
| *UIS8* (–448 to –460) | pTC12; GK2/GK41 and GK42/GK3 | AGGCGCTATTTTG → tGatcaatgcaTG | pGV33 and pGV34 |
| | pGV22[b]; GK2/GK37 and GK38/GK3 | Combine *uis6* and *uis7* mutations | pGV35 and pGV36 |
| | pGV34[b]; GK2/GK37 and GK38/GK3 | Combine *uis7* and *uis8* mutations | pGV37 and pGV38 |
| | pGV34; GK2/GK33 and GK34/GK3 | Combine *uis6* and *uis8* mutations | pGV39 and pGV40 |
| | pGV22; GK2/GK37 and pGV34; GK3/GK38 | Combine *uis6*, *uis7* and *uis8* mutations | pGV41 and pGV42 |
| Deletion mutations | | | |
| *DAL4* promoter | pTC12; GK2/GK7 | Delete 286 bp of the *DAL4* promoter | pGV45 |
| *DAL1* promoter | pTC12; GK2/GK8 | Delete 269 bp of the *DAL1* promoter | pGV46 |
| *DAL1* promoter | pTC12; GK2/GK20 | Delete 522 bp of the *DAL1* promoter | pGV47 |

[a] Buckholz and Cooper (1991)
[b] This work

probes were radioactively labeled by random priming (Roche Molecular Biochemicals). Standard prehybridization, hybridization, and washing conditions were followed (Ausubel et al. 1994).

## Results

### Steady-state *DAL1* and *DAL4* expression (mRNA) profiles

To establish the basic *DAL1* and *DAL4* expression profiles, we analyzed steady-state RNA from wild-type strain TCY1, grown in glucose-proline medium with and without OXLU. Although an 829-bp intergenic region is shared by *DAL1* and *DAL4*, the genes are regulated differently. *DAL4* is much more inducer-responsive than *DAL1* (Fig. 1), as reported earlier on the basis of enzyme activities in Σ1278b-based strains.

To identify the *cis*-acting elements that mediate transcription of the two genes, we constructed in-frame *DAL1*- and *DAL4-lacZ* fusions pGV1 and pGV2, respectively. Deletion of the *DAL4* third of the intergenic region (pGV46) had little effect on *DAL1* expression; and a larger deletion eliminated all *DAL1* expression (pGV47; Fig. 2). Deletion of the *DAL1* third (pGV45) decreased induced *DAL4* expression by 2/3 (Fig. 2). All *DAL1-lacZ* and *DAL4-lacZ* expression was highly nitrogen catabolite repression (NCR)-sensitive, Gln3p-dependent, and Dal81p-dependent (Fig. 2 and data not shown). Induced β-galactosidase production from *DAL1-lacZ* and *DAL4-lacZ* differed by 15-fold.

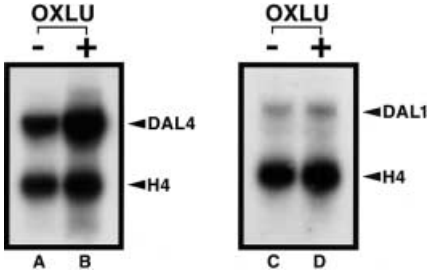


**Fig. 1** *Saccharomyces cerevisiae DAL1* and *DAL4* expression in the presence (*lanes B, D*) and absence (*lanes A, C*) of the allantoin pathway inducer oxalurate (*OXLU*). Poly(A)$^+$ RNA was prepared from wild-type strain TCY1 grown in glucose-proline yeast nitrogen base (YNB) medium with (+) or without (–) of 0.5 mM OXLU. Histone H4 served as a control for mRNA loadings and transfer efficiencies

### Contribution of $UAS_{NTR}$-homologous sequences to *DAL4* expression

To determine the contribution of individual intergenic GATA sequences to *DAL4* expression, each was destroyed by substitution mutations that did not otherwise alter the intergenic region. Transformants were assayed in both the presence and absence of inducer, because $UAS_{NTR}$ and $UIS_{ALL}$ are known to function synergistically in supporting inducer-mediated transcription (Yoo et al. 1989). The single mutant plasmids were also assayed in cells grown with asparagine as sole nitrogen source; and none of them supported reporter gene expression (data not shown). Activities observed in wild-type cells grown in inducer-free medium were too small in some cases to confidently compare alleles (Fig. 3). However, *DAL4* expression was highly inducible (fold induction) with all but two plasmids. The fold induction declined both with pGV14 (*gata2,3*), due to higher uninduced and lower induced levels, and with pGV16 (*gata4,5*), because expression was largely lost. Single *gata1* (pGV4), *gata2* (pGV6), and *gata3* (pGV8) mutations, while not significantly altering the fold induction, decreased the induced β-galactosidase levels incrementally by 1.6- to 3.4-fold; and *gata4* (pGV10) and *gata5* (pGV12) mutations resulted in somewhat greater decreases, 3.4- and 4.0-fold respectively (Fig. 3). The loss of induced *lacZ* expression in a *gata2,3* double mutant (pGV14) was roughly the same as seen in either single mutant, arguing that either

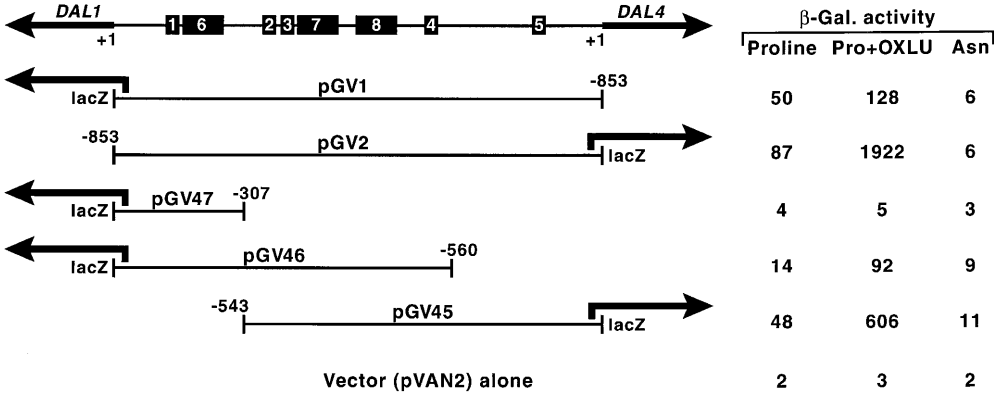


**Fig. 2** 5′ Deletion analysis of the *DAL1* and *DAL4* upstream regions. A schematic of the *DAL1–DAL4* intergenic region (*top*) indicates $UAS_{NTR}$ (*small black boxes, 1–5*) and $UIS_{ALL}$ (*large boxes, 6–8*) homologous sequences. Coordinates indicate the 5′ termini of the remaining promoter DNA. *Arrows marked lacZ* indicate the intergenic region fusion point with *lacZ*. Transformants (TCY1 recipient) were grown in YNB-proline (0.1%) medium without (*Proline*) and with (*Pro + OXLU*) 0.5 mM OXLU; 0.1% asparagine (*Asn*) was also used as a nitrogen source. β-Galactosidase (*β-Gal.*) activities are expressed in Miller units

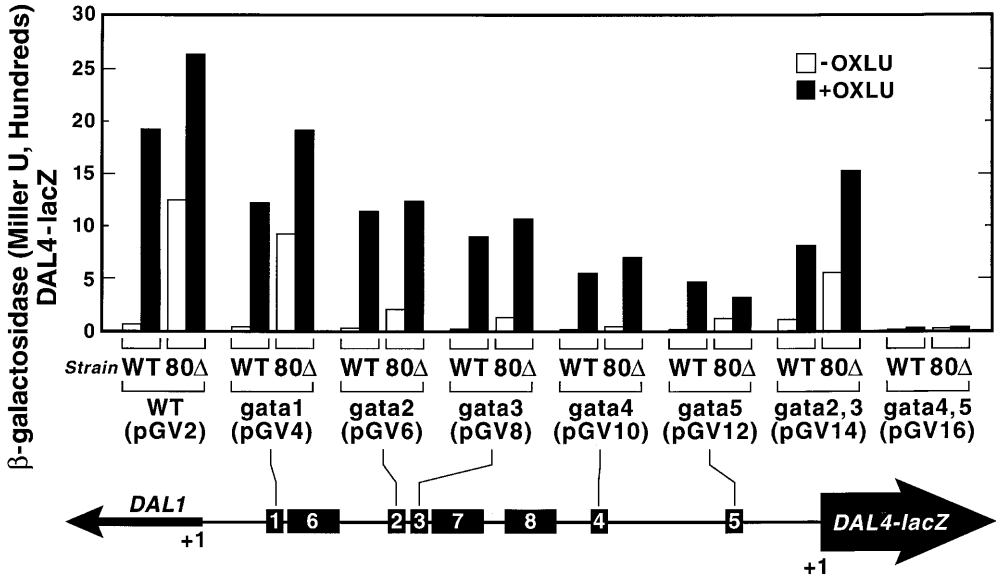


**Fig. 3** Single and combinational mutation analyses of putative *GATA* elements in the *DAL1–DAL4* intergenic region and their contribution to *DAL4-lacZ* expression. Elements were mutated in the context of the full-length intergenic region. *Small lettering* indicates the specific *GATA* element(s) mutated with the corresponding plasmid *in brackets below*. The native *DAL4-lacZ* fusion (*large arrow*) is indicated by *WT* (*pGV2*). These plasmids were transformed into strains TCY1 (*WT*) and TCY17 (*dal80*Δ). Transformants were grown in YNB-proline media in the absence (*open bars*) or presence (*solid bars*) of 0.5 mM OXLU

sequence would suffice (Fig. 3). In contrast, a *gata4,5* double mutant (pGV16) was synthetic, reducing *β*-galactosidase production to background levels (Fig. 3). While each GATA sequence contributes to overall induced *DAL4* expression, the pair of GATAs closest to the *DAL4* TATA elements are by far most crucial.

Evaluating the contributions of particular GATA sequences to *DAL1* and/or *DAL4* expression is complicated by the fact that such sequences are potential binding sites not only for Gln3p and Gat1p, but also for the repressor Dal80p (Cunningham et al. 1996). When this complication is eliminated by performing the experiment in a *dal80* mutant, uninduced expression levels can be easily compared. All but the *gata1* mutation decreased uninduced level expression 6- to 25-fold, arguing that GATAs 2–5 all functioned (Fig. 3). Somewhat surprisingly, the *gata2,3* double mutation caused a significantly smaller decrease in the absence of inducer than either of the corresponding single mutations (pGV14, pGV6, pGV8). In the presence of inducer, the *gata4* (pGV10) and *gata5* (pGV12) mutations possessed the strongest phenotypes; and the *gata4,5* double mutation (pGV16) totally destroyed *lacZ* expression (Fig. 3).

## Contribution of $UIS_{ALL}$-homologous sequences to *DAL4* expression

In addition to the five GATA sequences, the *DAL1–DAL4* intergenic region contains three $UIS_{ALL}$-homologous sequences, potential binding sites for the Dal82p that is required for inducer-dependent transcription. Mutating individual $UIS_{ALL}$-homologous sequences [*UIS6* (pGV22) and *UIS8* (pGV34)] reduced induced *DAL4-lacZ* expression 1.7- and 2.5-fold, respectively; and the *UIS7* mutation (pGV28) was without effect (Fig. 4). In contrast to expectation, uninduced *DAL4-lacZ* expression increased 6.5-fold in a *uis7* mutant relative to the wild type (Fig. 4). Double and triple *uis* mutations produced stronger phenotypes. The *uis6,7,8* triple mutation (pGV42) supported the least *β*-galactosidase production, which was also inducer-independent (Fig. 4). Any double mutant containing a *uis8* mutation [*uis7,8* (pGV38) and *uis6,8* (pGV40)] supported a similarly low induced *lacZ* expression, which was less than in the *uis6,7* mutant (pGV36). Any double mutant containing a *uis7* mutation exhibited significantly more expression in the absence of inducer (Fig. 4).

While all three $UIS_{ALL}$-homologous sequences contributed to overall inducer-responsive *DAL4* expression, their contributions were not equal; and in one case, the element had the opposite function to the others. Inducer-responsiveness of the *DAL* genes depends upon Dal82p, which binds to $UIS_{ALL}$ elements (Dorrington and Cooper 1993). Therefore, as a control, we compared results obtained in a wild type with those in *dal82*Δ. Expression in *dal82*Δ strain SS400 was unaffected by inducer (Fig. 4). However, in every case, the uninduced expression level was greater than that in the wild type, further substantiating that one or more of the $UIS_{ALL}$-homologous sequences could behave as a negative regulator when inducer is absent.

## Contribution of $UAS_{NTR}$-homologous sequences to *DAL1* expression

We similarly evaluated contributions of the five $UAS_{NTR}$-homologous sequences to *DAL1* expression. In

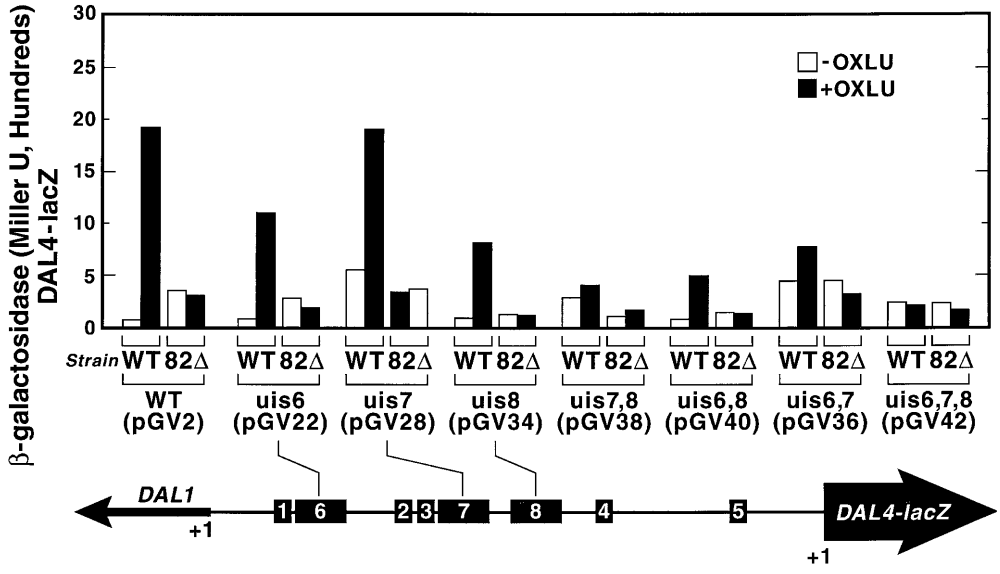


**Fig. 4** Single and combinational mutation analyses of putative $UIS_{ALL}$ elements present in the *DAL1–DAL4* intergenic region and their contribution to *DAL4-lacZ* expression. $UIS_{ALL}$-homologous elements were mutated in the context of the full-length intergenic region. *Small lettering* indicates the specific $UIS_{ALL}$ element(s) mutated with the corresponding plasmid *in brackets below*. The native *DAL4-lacZ* fusion is indicated by *WT* (*pGV2*). Plasmids were transformed into strain TCY1 (*WT*) or SS400 (*dal82*Δ). Transformants were grown in YNB-proline medium in the absence (*open bars*) or presence (*solid bars*) of 0.5 mM OXLU

contrast to *DAL4, DAL1* expression was much less inducible (less than 2.5-fold) and mutating the GATA sequence most proximal to *DAL1* (*gata1*; pGV3) had no significant effect on the expression observed in a wild-type strain (Fig. 5). Mutating each of the remaining four GATAs decreased the induced *DAL1-lacZ* expression less than 2.0-fold; and similar results occurred with *gata2,3* and *gata4,5* double-mutant plasmids.

Stronger mutant phenotypes occurred when the analysis was performed in a *dal80* mutant background. The first and most striking characteristic of the data is seen by comparing *DAL4-lacZ* and *DAL1-lacZ* expression. For *DAL4-lacZ*, uninduced β-galactosidase production increases in *dal80* mutants, but never exceeds the levels seen in an induced strain containing a wild-type plasmid (Fig. 3, pGV2). In contrast, uninduced *DAL1-lacZ* expression in a *dal80* strain background increases to a much higher level than induced *DAL1-lacZ* expression in both wild-type and *dal80* mutant strains (Fig. 5, pGV1). This profile is characteristic of a NCR-sensitive gene whose expression is largely inducer-independent, e.g., *DAL5* (Cooper 1996).

In the *dal80* mutant background, the differing contributions of the GATA-homologous sequences to overall *DAL1* expression are more apparent. Mutation of *GATA2*, *3*, or *4* decreased the uninduced expression 4.8- to 8.0-fold relative to the wild type (Fig. 5, pGV5, pGV7, pGV9). However, *lacZ* expression in a *gata5* mutant (pGV11) decreases less than 2.0-fold relative to the wild type (pGV1), arguing that its contribution to *DAL1* expression is quite limited. In contrast, mutating the *GATA1* sequence (pGV3) increases the amount of *lacZ* expression observed in a *dal80* mutant significantly above the level supported by the wild-type pGV1.

## Contribution of $UIS_{ALL}$-homologous sequences to *DAL1* expression

Induced β-galactosidase production, from the wild-type *DAL1-lacZ* pGV1, was about 2.5-fold greater than the uninduced level (Fig. 6). The only effect of mutating the UIS most proximal to *DAL1* (*UIS6*) was a 3.0-fold decrease in uninduced reporter expression (Fig. 6, pGV21). Mutating the UIS most proximal to *DAL4* (*UIS8*) decreased the induced β-galactosidase production from *DAL1-lacZ*, but only very modestly; and it did not affect the uninduced levels (Fig. 6, pGV33). The *uis6,8* double mutant (pGV39) exhibited uninduced expression similar to that supported by the *uis6* mutation and an induced level more similar to that seen with the *uis8* mutation (Fig. 6).

Mutating the central $UIS_{ALL}$-homologous sequence, *UIS7*, generated the strongest and also most surprising phenotype (Fig. 6, pGV27). Uninduced and induced *lacZ* expression increased 5.0- and 2.0-fold, respectively. In other words, the *UIS7* sequence behaved more as a negative than a positive regulator of *DAL1* expression. An analogous response has been reported for an inducer-responsive element situated upstream of *CAR1* (Kovari et al. 1990). The *uis6,7* double mutant exhibited a phenotype similar to that of a *uis7* single mutant, with respect to the wild type. In contrast, a *uis7,8* double mutant exhibited a phenotype that more closely resembled the *uis8* single-mutant phenotype. Together, these data suggest the *UIS7* sequence down-regulates *DAL1* expression, supported in part by *UIS8*. It must be emphasized, however, that the negative regulation is stronger than the positive. When the experiment was

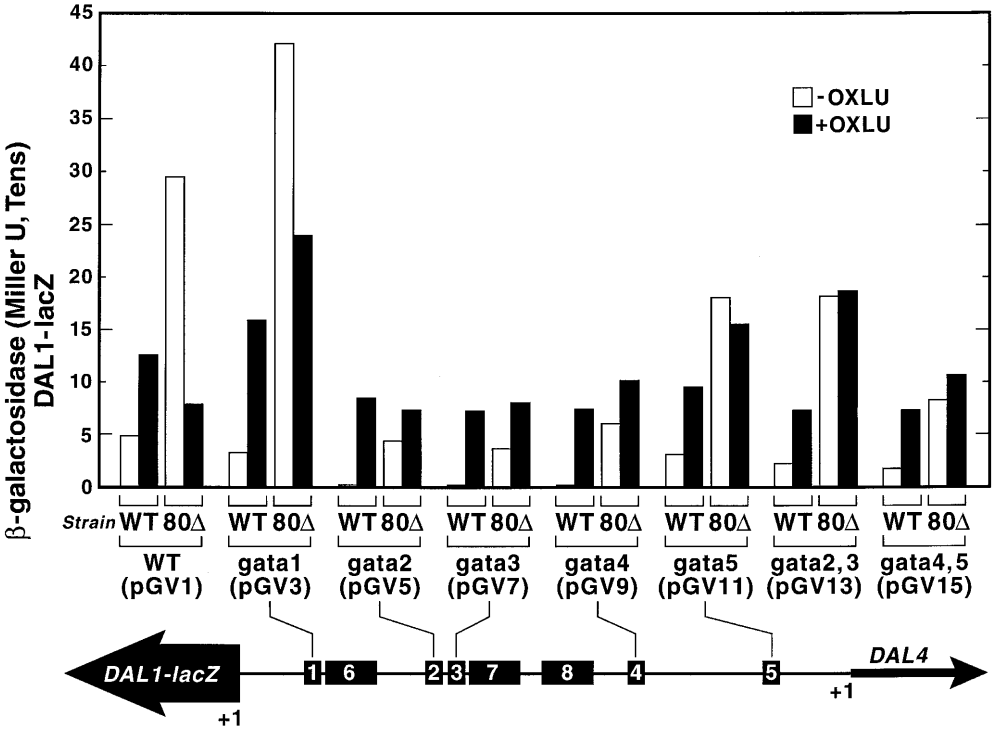


**Fig. 5** Single and combinational mutation analyses of putative *GATA* elements in the *DAL1–DAL4* intergenic region and their contribution to *DAL1-lacZ* expression. The experiment was performed as described in Fig. 3, except that *lacZ* was fused to the *DAL1* end of the intergenic region (*large arrow*)

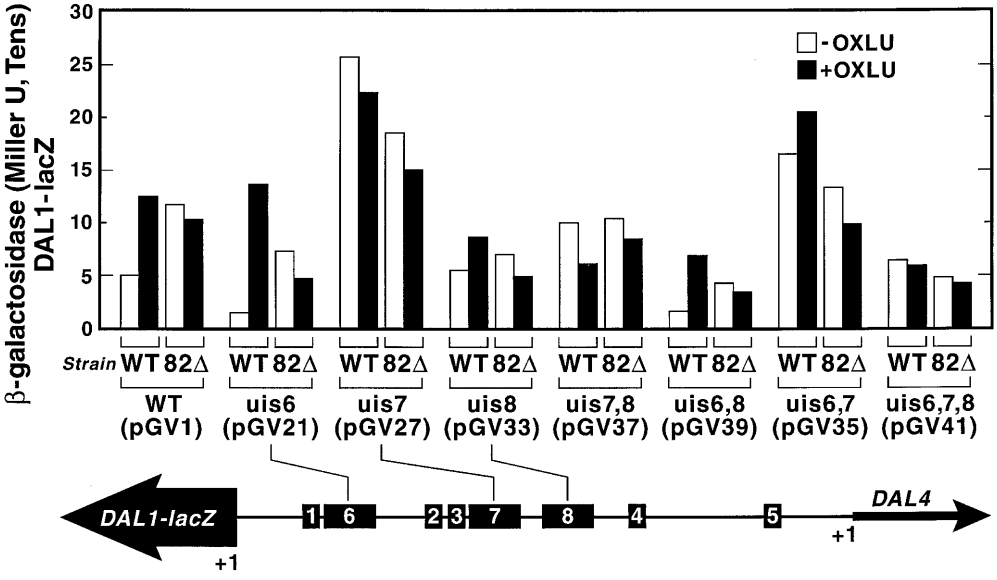


**Fig. 6** Single and combinational mutation analyses of putative $UIS_{ALL}$ elements in the *DAL1–DAL4* intergenic region and their contribution to *DAL1-lacZ* expression. The experiment was performed as described in Fig. 4, except that *lacZ* was fused to the *DAL1* end of the intergenic region (*large arrow*)

repeated in a *dal82*Δ, the major effect observed was a loss of inducer responsiveness, i.e., all of the uninduced *lacZ* levels were as high or higher than the induced levels (Fig. 6).

## Discussion

The above data assess the contributions to *DAL1*- and *DAL4-lacZ* expression of each $UAS_{NTR}$- or $UIS_{ALL}$-homologous sequence in the *DAL1–DAL4* intergenic region. In a wild-type background, in the presence of inducer, all five GATAs appear to participate in *DAL4-lacZ* expression with *GATA2–GATA3* and *GATA4–GATA5* making relatively equal contributions. The *gata2,3* and *gata4,5* double mutants, however, argue that *GATA4* and *GATA5* are the most important. Similar conclusions are drawn from the experiments in a *dal80* background grown with inducer. Without inducer, however, only GATAs2–5 appear to function.

Single mutation data argue that induced *DAL4* expression depends somewhat more heavily on *UIS8* than

*UIS6*, while *UIS7* is not a positive participant in *DAL4* induction. Data with double and triple mutations, however, suggest that *UIS7* can play a limited positive role, if either *UIS8* and *UIS6* is mutated (Fig. 4, pGV40, pGV42). In contrast with expectation, *UIS7* appears to play a negative role in the absence of inducer.

From previous detailed studies of the *CAR1*, *CAR2*, *DAL5*, *DAL7*, and *DUR1,2* promoters, one would have predicted that *GATA4*, *GATA5*, and *UIS8* were the most likely participants in *DAL4-lacZ* expression, based both on their sequences and on their orientation and location. (Genbauffe, El Berry, Daugherty, and Cooper, unpublished data; Kovari et al. 1990; Park et al. 1999; Rai et al. 1989, 1999). Here data and expectation are similar, but not congruent. However, previous literature also reports that a DNA fragment carrying *UIS7* is a much better competitor for a standard Dal82p-binding DNA fragment than *UIS8* (Dorrington and Cooper 1993). From this, *UIS7* would have been hypothesized to be the more likely candidate as the *cis*-acting element mediating induced *DAL4* expression. Data presented here are consistent with both *UIS7* and *UIS8* participating in gene function but in opposite ways which is not predictable a priori.

For *DAL1-lacZ* expression, predictions of the participating *cis*-acting elements are less easily made, but the most reasonable choices are *GATA1*, *GATA2*, and *GATA3*. UIS elements cannot be considered significant, because *DAL1* expression is so little affected by the addition of inducer. From data with single intergenic mutations in a wild-type background, *GATA2–4* are about equal participants in the absence of inducer; and the participation of *GATA5* is marginal. However, from double mutants, uninduced *DAL1-lacZ* expression can proceed quite well in the absence of both *GATA2* and *GATA3*. Mutation of *GATA1* yields a slight increase in the induced level of expression, whereas in *gata2*, *gata3*, and *gata4* mutants, the induced expression is slightly depressed. Although most of the conclusions drawn from the wild type are substantiated by data in the *dal80* background, there is a marked change. The wild-type plasmid (pGV1) exhibits greater expression in the absence of inducer than in its presence; and *GATA1* is responsible for this pattern of expression.

Expectation and observation differ even more for the case of *UIS* participation in *DAL1* expression. First and foremost, since *DAL1* expression is not particularly inducer responsive (Fig. 1), $UIS_{ALL}$ participation is not expected. However, detailed analysis reveals that UIS elements definitely participate in *DAL1* transcription. Second, to the extent that *DAL1* is OXLU-responsive, one would predict *UIS6* and *UIS7* to most likely account for any induction observed. *UIS6*, the $UIS_{ALL}$-homologous sequence most proximal to *DAL1*, does not appear to participate; and, surprisingly, *UIS7* appears to be a negative rather than positive regulator . The putative element least expected to participate in *DAL1* expression, *UIS8*, is the most necessary. Considering single- and double-mutant data together, *UIS8* continues to be most necessary.

The appearance of *UIS7* as a putative, negatively acting element deserves further comment as, at face value, it seems to contradict much of the literature concerning $UIS_{ALL}$ and Dal82p. A similar phenomenon was observed during dissection of the *CAR1* promoter (Kovari et al. 1990). In that instance, placing an arginine-dependent UAS element downstream of a constitutively acting UAS, $UAS_{C2}$, resulted in a six-fold decrease in reporter gene expression when glutamate was provided as sole nitrogen source. This "negative regulation" disappeared, however, when arginine was used in place of glutamate as the nitrogen source (Fig. 7, pLK78, pLK105 in Kovari et al. 1990). The explanation offered for *CAR1* was that if a protein that is not functioning in transcription, i.e., serving as a UAS due to the absence of the inducer (arginine) upon which its operation depends, binds to a site downstream of a functioning UAS ($UAS_{C2}$), it will repress transcription supported by it. We suggest that *UIS7* is behaving similarly. *UIS7* is not functioning as a UIS element as far as *DAL1* expression is concerned and hence behaves like a negatively acting element, because it is situated between the *cis*-acting elements responsible for *DAL1* transcription and the TATA element.

This work focused only on known allantoin-pathway, *cis*-acting elements. The *DAL1–DAL4* intergenic region may well contain other *cis*-acting elements that have gone unnoticed, but are important to the operation of the elements we analyzed. That such additional elements also participate in the expression of allantoin and arginine pathway genes is well documented (Dubois and Messenguy 1997; Park et al. 1999; Rai et al. 1999; Smart et al. 1996). The presence of such elements would only complicate a straightforward correlative analysis yet further. An example of this may be seen with respect to the *DAL1* and *DAL4* TATA elements. *DAL4* is expressed much more strongly than *DAL1*. This correlates with the presence of three potentially strong TATA sequences, TATAAA, TATATA, and TATAT immediately upstream of *DAL4*, whereas only two such sequences, TATAG and TATAT, appear upstream of *DAL1*.

Beyond identifying the relative contributions of the *DAL1–DAL4 cis*-acting sequences, these data demonstrate that:

1. The actual participation of various $UAS_{NTR}$- and $UIS_{ALL}$-homologous sequences in *DAL1*- and *DAL4-lacZ* expression could only be predicted with quite limited success, even though they were based on previously reported, detailed analyses of five highly related promoters.
2. *DAL1* and *DAL4* do not share a set of *cis*-acting elements equally; and, further, elements for the expression of the two genes cannot be predicted from their locations.
3. Even shared elements do not always function in the same way for the two genes.

4. The potential participation of a given element and its contribution to *DAL1* and *DAL4* expression is a function not only of the element and its location, but also of the other elements functioning in proximity to it.
5. The $UIS_{ALL}$-homologous sequence predicted, on the basis of in vitro DNA-binding experiments (Dorrington and Cooper 1993), as being the most likely to be responsible for induced *DAL4* expression did not contribute to induction.

The *DAL1* expression profile also demonstrates that the presence of both $UAS_{NTR}$- and $UIS_{ALL}$-homologous sequences in a gene's promoter region is not necessarily indicative of inducibility as expected a priori. Therefore, correlations of genome-wide expression profiles with the presence of sequences homologous to known transcription factor-binding sites or found repeated within the promoter sequences of co-regulated genes may be a legitimate starting point for investigating their potential function. However, with our current state of technology, they cannot be considered as an end-point with which to draw rigorous conclusions.

**Acknowledgements** We thank Tim Higgins for preparing the artwork, and the UBC Wine Research Center and UT Yeast Groups for suggested improvements to the manuscript. This work was supported by NSERC grant 217271-99 to H.J.J.V. and NIH grant GM-35642 to T.G.C.

## References

Angermayr M, Bandlow W (1997) The type of basal promoter determines the regulated or constitutive mode of transcription in the common control region of the yeast gene pair *CYC1*/*RIO1*. J Biol Chem 272:31630–31635

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1994) Current protocols in molecular biology. Wiley, New York

Bell PJ, Bissinger PH, Evans RJ, Dawes IW (1995) A two-reporter gene system for the analysis of bi-directional transcription from divergent *MAL6T-MAL6S* promoter in *Saccharomyces cerevisiae*. Curr Genet 28:441–446

Bell PJ, Higgins VJ, Dawes IW, Bissinger PH (1997) Tandemly repeated 147 bp elements cause structural and functional variation in divergent *MAL* promoters of *Saccharomyces cerevisiae*. Yeast 13:1135–1144

Blinder D, Magasanik B (1995) Recognition of nitrogen-responsive upstream activation sequences of *Saccharomyces cerevisiae* by the product of the *GLN3* gene. J Bacteriol 177:4190–4193

Buckholz RG, Cooper TG (1991) The allantoinase (*DAL1*) gene of *Saccharomyces cerevisiae*. Yeast 7:913–923

Bysani N, Daugherty JR, Cooper TG (1991) Saturation mutagenesis of the $UAS_{NTR}$ (GATAA) responsible for nitrogen catabolite repression-sensitive transcriptional activation of the allantoin pathway genes in *Saccharomyces cerevisiae*. J Bacteriol 173:4977–4982

Cooper TG (1996) Regulation of allantoin catabolism in *Saccharomyces cerevisiae*. Mycota 3:139–169

Cunningham TS, Cooper TG (1993) The *Saccharomyces cerevisiae* DAL80 repressor protein binds to multiple copies of GATAA-containing sequences ($URS_{GATA}$). J Bacteriol 175:5851–5861

Cunningham TS, Dorrington RA, Cooper TG (1994) The *UGA4* $UAS_{NTR}$ site required for GLN3-dependent transcriptional activation also mediates DAL80-responsive regulation and DAL80 protein binding in *Saccharomyces cerevisiae*. J Bacteriol 176:4718–4725

Cunningham TS, Svetlov VV, Rai R, Smart W, Cooper TG (1996) Gln3p is capable of binding to $UAS_{NTR}$ elements and activating transcription in *Saccharomyces cerevisiae*. J Bacteriol 178:3470–3479

Dorrington RA, Cooper TG (1993) The DAL82 protein of *Saccharomyces cerevisiae* binds to the *DAL* upstream induction sequence (UIS). Nucleic Acids Res 21:3777–3784

Dubois E, Messenguy F (1997) Integration of the multiple controls relating the expression of the arginase gene *CAR1* of *Saccharomyces cerevisiae* in response to different nitrogen signals: role of Gln3p, ArgRp-Mcm1p, and Ume6p. Mol Gen Genet 253:568–580

Friesen H, Hepworth SR, Segall J (1997) An Ssn6-Tup1-dependent negative regulatory element controls sporulation-specific expression of *DIT1* and *DIT2* in *Saccharomyces cerevisiae*. Mol Cell Biol 17:123–134

Geitz D, St. Jean A, Woods RA, Schiestl RH (1992) Improved method for high-efficiency transformation of intact yeast cells. Nucleic Acids Res 20:1425

Hahn S, Pinkham J, Wei R, Miller R, Guarente L (1988) The *HAP3* regulatory locus of *Saccharomyces cerevisiae* encodes divergent overlapping transcripts. Mol Cell Biol 8:655–663

Halfter H, Muller U, Winnacker EL, Gallwitz D (1989) Isolation and DNA-binding characteristics of a protein involved in transcription activation of two divergently transcribed, essential yeast genes. EMBO J 8:3029–3037

Hofman-Bang J (1999) Nitrogen catabolite repression in *Saccharomyces cerevisiae*. Mol Biotech 12:35–73

Inoue H, Nojima H, Okayama H (1990) High efficiency transformation of *Escherichia coli* with plasmids. Gene 96:23–28

Johnston M, Davis RW (1984) Sequences that regulate the divergent *GAL1–GAL10* promoter in *Saccharomyces cerevisiae*. Mol Cell Biol 4:1440–1448

Kovari L, Sumrada R, Kovari I, Cooper TG (1990) Multiple positive and negative *cis*-acting elements mediate induced arginase (*CAR1*) gene expression in *Saccharomyces cerevisiae*. Mol Cell Biol 10:5087–5097

Kraakman LS, Mager WH, Maurer KT, Nieuwint RT, Planta RJ (1989) The divergently transcribed genes encoding yeast ribosomal proteins L46 and S24 are activated by shared RPG-boxes. Nucleic Acids Res 17:9693–9706

Kruglyak S, Tang H (2000) Regulation of adjacent yeast genes. Trends Genet 16:109–111

Liu Y, Xiao W (1997) Bidirectional regulation of two DNA-damage-inducible genes, *MAG1* and *DDI1*, from *Saccharomyces cerevisiae*. Mol Microbiol 23:777–789

Lyons TJ, Gasch AP, Gaither LA, Botstein D, Brown PO, Eide DJ (2000) Genome-wide characterization of the Zap1p zinc-responsive regulon in yeast. Proc Natl Acad Sci USA 97:7957–7962

Miller JH (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Osley MA, Gould J, Kim S, Kane MY, Hereford L (1986) Identification of sequences in a yeast histone promoter involved in periodic transcription. Cell 45:537–544

Park H-D, Luche RM, Cooper TG (1992) The yeast *UME6* gene product is required for transcriptional repression mediated by the *CAR1 URS1* repressor binding site. Nucleic Acids Res 20:1909–1915

Park H-D, Scott S, Rai R, Dorrington R, Cooper TG (1999) Synergistic operation of the *CAR2* (ornithine transaminase) promoter elements in *Saccharomyces cerevisiae*. J Bacteriol 181:7052–7064

Rai R, Genbauffe FS, Sumrada RA, Cooper TG (1989) Identification of sequences responsible for transcriptional activation of the allantoate permease gene in *Saccharomyces cerevisiae*. Mol Cell Biol 9:602–608

Rai R, Daugherty JR, Cunningham TS, Cooper TG (1999) Overlapping positive and negative GATA factor binding sites

mediate inducible *DAL7* gene expression in *Saccharomyces cerevisiae*. J Biol Chem 274:28026–28034
Schlapp T, Rodel G (1990) Transcription of two divergently transcribed yeast genes initiates at a common oligo(dA-DT) tract. Mol Gen Genet 223:438–442
Schure EG ter, Riel NAW van, Verrips CT (2000) The role of ammonia metabolism in nitrogen catabolite repression in *Saccharomyces cerevisiae*. FEMS Microbiol Rev 24:67–83
Siliciano PG, Tatchell K (1984) Transcription and regulatory signals at the mating type locus in yeast. Cell 37:969–978
Smart WC, Coffman JA, Cooper TG (1996) Combinatorial regulation of the *Saccharomyces cerevisiae CAR1* (arginase) promoter in response to multiple environmental signals. Mol Cell Biol 16:5876–5887
Struhl K (1985) Nucleotide sequence and transcriptional mapping of the yeast *pet56-his3-ded1* gene region. Nucleic Acids Res 13:8587–8601
Thuriaux P, Mariotte S, Buhler JM, Sentenac A, Vu L, Lee BS, Nomura M (1995) Gene *RPA43* in *Saccharomyces cerevisiae* encodes and essential subunit of RNA polymerase I. J Biol Chem 270:24252–24257
Van Vuuren HJJ, Daugherty JR, Rai R, Cooper TG (1991) Upstream induction sequence, the *cis*-acting element required for response to the allantoin pathway inducer and enhancement of operation of the nitrogen-regulated upstream activation sequence in *Saccharomyces cerevisiae*. J Bacteriol 173:7186–7195
Viljoen M, Volschenk H, Young RA, Van Vuuren HJJ (1999) Transcriptional regulation of the *Schizosaccharomyces pombe* malic enzyme gene, *mae2*. J Biol Chem 274:9969–9975
West RW, Yocum RR, Ptashne M (1984) *Saccharomyces cerevisiae GAL1–GAL10* divergent promoter region: location and function of the upstream activating sequence *UASG*. Mol Cell Biol 4:2467–2478
Wiame J-M, Grenson M, Arst H (1985) Nitrogen catabolite repression in yeasts and filamentous fungi. Adv Microb Physiol 26:1–87
Yoo HS, Cooper TG (1989) The *DAL7* promoter consists of multiple elements that cooperatively mediate regulation of the gene's expression. Mol Cell Biol 9:3231–3243