Molecular and Cellular Biology, Oct. 1984, p. 1985–1998 Vol. 4, No. 10
0270-7306/84/101985-14$02.00/0
Copyright © 1984, American Society for Microbiology

Use of *lacZ* Fusions to Delimit Regulatory Elements of the Inducible Divergent *GAL1-GAL10* Promoter in *Saccharomyces cerevisiae*

R. ROGERS YOCUM,[1,2]* SUSAN HANLEY,[2] ROBERT WEST, JR.,[1] AND MARK PTASHNE[1]

The Biochemical Laboratories, Harvard University, Cambridge, Massachusetts 02138,[1] and Biotechnica International, Cambridge, Massachusetts 02140[2]

Received 10 March 1984/Accepted 2 July 1984

We present the DNA sequence of a 914-base pair fragment from *Saccharomyces cerevisiae* that contains the *GAL1-GAL10* divergent promoter, 140 base pairs of *GAL10* coding sequence, and 87 base pairs of *GAL1* coding sequence. From this fragment, we constructed four pairs of *GAL1-lacZ* and *GAL10-lacZ* fusions on various types of yeast plasmid vectors. On each type of vector, the fused genes were induced by galactose and repressed by glucose. The response of a *GAL1-lacZ* fusion to *gal4* and *gal80* regulatory mutations was similar to the response of intact chromosomal *GAL1* and *GAL10* genes. A set of deletions that removed various portions of the *GAL10* regulatory sequences from a *GAL10-CYC1-lacZ* fusion was constructed in vitro. These deletions defined a relatively guanine-cytosine-rich region of 45 base pairs that contained sequences necessary for full-strength galactose induction and an adjacent guanine-cytosine rich 55 base pairs that contained sequences sufficient for weak induction.

The yeast *Saccharomyces cerevisiae* can grow on galactose as a sole carbon source. Galactose is transported into the cell by a specific galactose permease and is converted to glucose-1-phosphate by the sequential action of three enzymes, galactokinase, α-D-galactose-1-phosphate uridyltransferase, and uridine diphosphoglucose-4-epimerase (6, 32). These three enzymes are encoded by a tightly linked cluster of genes named *GAL1*, *GAL7*, and *GAL10*, respectively (10). If cells are grown on glucose or glycerol and then switched to galactose, *GAL1*, *GAL7*, and *GAL10* are coordinately induced at least 1,000-fold at the level of transcription (11, 25, 59–61). This dramatic example of regulation in a genetically tractable organism provides an ideal system for studying a eucaryotic gene control mechanism.

Expression of *GAL1*, *-7*, and *-10* is governed by at least two distinct regulatory circuits, a galactose-specific induction system (11) and a more general glucose repression system (1, 37). Much is known about galactose-specific induction (see reference 47 for review). Classical genetics and recent molecular studies have led to the following model. In the absence of galactose, *GAL1*, *GAL7*, and *GAL10* are repressed by a negative regulatory protein encoded by the *GAL80* gene (11). Recessive *gal80* mutations cause constitutive expression of *GAL1*, *GAL7*, and *GAL10*, whereas dominant *GAL80*^{s} mutations confer an uninducible phenotype (11–13, 45). A positive regulatory protein, encoded by the *GAL4* gene, is required for expression of *GAL1*, *GAL7*, and *GAL10*, even in the absence of wild-type *GAL80* (11–13, 25, 35, 36). Dominant *GAL4*^{c} mutations cause constitutive expression of *GAL1*, *GAL7*, and *GAL10*, whereas recessive *gal4* mutations confer an uninducible phenotype (11, 35, 45). Both the *GAL4* and *GAL80* proteins are constitutively produced in low amounts (24, 28, 36 48). Complementation analysis of various pairs of *GAL4*^{c} and *GAL80*^{s} alleles has suggested that the *GAL4* and *GAL80* gene products interact with each other (45). It is thought that the inducer (most likely galactose) binds to a site on the *GAL80* protein resulting in release of the *GAL4* protein. The free *GAL4* protein then acts to turn on transcription of *GAL1*, *GAL7*, and *GAL10*, presumably by binding to DNA sites near the 5′ ends of the genes (24, 25, 35, 36, 45, 48).

The *GAL4* gene has been cloned (24, 28, 31, 31a). Overproduction of the *GAL4* protein is sufficient to induce *GAL1*, *GAL7*, and *GAL10*, even in a glucose medium (24, 28). However, the details of the mechanism of action of *GAL4* and *GAL80* are still unknown.

Two other genes, *GAL3* and *GAL11*, are necessary for a wild-type *GAL* phenotype. *gal3* mutants, or "slow adapters," take over 20 h instead of the normal 10 min for the rate of transcription of *GAL1*, *GAL7*, and *GAL10* to reach steady-state levels after galactose induction (4, 13, 29). *gal11* mutants produce only 20% of wild-type levels of the three Leloir enzymes after induction with galactose (43). The actual functions of *GAL3* and *GAL11* are still obscure. The regulatory loci, *GAL3*, *GAL4*, *GAL11*, and *GAL80*, segregate independently from each other as well as from the *GAL1,7,10* cluster.

Little is known about the genetics or biochemistry of glucose repression. It is epistatic to galactose induction; growth in glucose plus galactose results in little or no expression of *GAL1*, *GAL7*, and *GAL10* (37). Some components of the glucose repression system, such as the products of the *GAL82* and *GAL83* genes, are specific for the *GAL1*, *-7*, and *-10* genes (37, 39). Other components, for example, the product of the *SNF1* gene, also act on genes coding for enzymes which utilize a variety of other carbon sources such as sucrose and maltose (5, 17). Glucose repression in *S. cerevisiae* is thought not to be mediated by intracellular levels of cyclic AMP as it is in *Escherichia coli* (38).

The *GAL1,7,10* gene cluster has been cloned and transcriptionally mapped (7, 51, 54, 59–61). None of these genes contains an intervening sequence. *GAL1* and *GAL10* are divergently transcribed from an intergenic region of 680 base pairs and *GAL7* is independently transcribed from a region of 650 base pairs between the end of *GAL10* and the

* Corresponding author.

beginning of *GAL7* (25, 44, 59–61). A 365-base pair fragment of DNA from the middle of the *GAL1-GAL10* intergenic region contains all signals necessary for galactose-induced transcription and glucose repression of a nearby transcription start site (23). The site within this fragment that is responsible for induction has been termed the "UAS," for upstream activation site (23).

We sought to learn more about the control of expression of *GAL1*, *GAL7*, and *GAL10* at the molecular level. We assumed that the *GAL* genes, like other yeast genes coding for proteins, are regulated through DNA sequences located near the 5' end of the coding sequences (2, 9, 16, 21, 49, 55, 62). First, we determined the DNA sequence of the *GAL1-GAL10* intergenic region. Next, we constructed gene fusions on plasmids in vitro between the *E. coli lacZ* gene and both *GAL1* and *GAL10* to provide an easy assay for expression of these genes (22, 49). The gene fusions were then shown to be regulated identically to the intact chromosomal *GAL1* and *GAL10* genes, in response to both different carbon sources and known regulatory mutations. Finally, a set of deletions was constructed in vitro which further delimited the DNA sequences involved in transcriptional control of *GAL10*.

MATERIALS AND METHODS

Plasmids were maintained and produced in *E. coli* MM294 (F^- *endA hsdR supE44 thiA*) or P90C [F^- Δ(*lac-pro*)XIII] (20). DNA sequences were determined by the chemical cleavage method (40). Manipulation of DNA sequences and homology searches were aided by the Stanford University SUMEX-<MOLGEN>SEQ computer program. All restriction enzymes, T4 DNA ligase, and the Klenow fragment of DNA polymerase I were purchased from New England Biolabs. Synthetic DNA linkers were from Collaborative Research or New England Biolabs, and 4'-bromo-5'-chloro-3'-indolyl-β-D-galactoside was from Bachem.

Plasmid constructions, yeast strain constructions, and yeast transformations were by standard methods (34, 56). Integrating plasmids were directed to a particular chromosomal location by linearizing the plasmid before transformation at a unique site in the sequences homologous to the target (46). Transformants containing a single copy of pRY171, -173, -181, or -183 (see below) were identified by digesting genomic DNA with *Bam*HI, which cleaves each of these plasmids only once, and screening for those which did not produce unit-length linear plasmid DNA (46).

All plasmids constructed for this study were derived from five plasmids supplied by others. pSc4816 contains a 914-base pair *Eco*RI-to-*Ava*I fragment which covers the region between *GAL1* and *GAL10* cloned with an *Eco*RI linker into the *Eco*RI site of pBR325 (T. St. John, unpublished data). pLG669 is a derivative of YEp24 (3) which contains *URA3*, a 2 μm replicator, and a *CYC1-lacI-lacZ* fusion (22). pLGSD5 was derived from pLG669 by replacing *CYC1* regulatory sequences with a 365-base pair *Dde*I-to-*Sau*3A fragment from the *GAL1-GAL10* divergent promoter region (23). pSZ211, a derivative of YIp32 (3), contains *LEU2*, *ARS1*, and a truncated, unfused *lacZ* gene (J. Szostak, personal communication). YCp19 contains *CEN4*, the centromere from yeast chromosome IV (58).

Deletions of pLGSD5 were constructed by cleaving the unique *Xho*I site at the *GAL10* end of the *GAL* UAS insert, digesting for various times with BAL31 exonuclease, ligating with *Xho*I linkers, and cutting with *Sac*I plus *Xho*I (34, 41). The plasmid backbone was purified by preparative gel electrophoresis (34) and ligated to the short *Xho*I-to-*Sac*I fragment of pLGSD5 that had not been treated with exonuclease. This resulted in a set of unidirectional deletions beginning at the *Xho*I site and extending into the *GAL* sequences.

The *S. cerevisiae* strains used in this study are listed in Table 1.

TABLE 1. *S. cerevisiae* strains

Strain	Genotype	Source
DBY745	α *ade1-100 leu2-3 leu2-112 ura3-52*	D. Botstein
JG50	**a** *gal4-3 leu2-3 leu2-112 trp1 his1*	J. Yarger
JG80	**a** *GAL80*^{s}*-96 trp1 his1*	J. Yarger
JG129	**a** *gal80*[a] *leu1 trp1*	J. Yarger
JG130	**a** *GAL4*^{c}*-c1 MEL1 ade his trp1*	J. Yarger
SHC22c	**a** *gal4-3 ade1-100 leu2-3 leu2-112 trp1 ura3-52*	DBY745 × JG50
SHE2b	**a** *gal80*[a] *ade1-100 leu1 ura3-52*	DBY745 × JG129
SHI6a	α *GAL4*^{c}*-c1 ade leu2-3 leu2-112 trp1 ura3-52*	DBY745 × JG130
SHJ1d	α *GAL80*^{s}*-96 his1 leu2-3 leu2-112 trp1 ura3-52*	DBY745 × JG80

[a] The particular *gal80* allele in these strains was not known.

Minimal selective media contained, per liter, 7 g of yeast nitrogen base (Difco Laboratories); 0.7 g of uracil dropout mix or leucine dropout mix (56) and 2% glucose (MinGlu); 2% glucose plus 2% galactose (MinGluGal); 2% galactose (MinGal); 2% galactose, 3% glycerol, plus 2% ethanol (MinGalGlyEth); or 3% glycerol plus 2% ethanol (MinGlyEth). Solid media contained, in addition, 2% agar. Indicator plates contained, in addition to the above ingredients, 40 mg of 4'-bromo-5'-chlor-3'-indolyl-β-D-galactoside per liter and 70 mM potassium phosphate buffer, pH 7.0, both added after autoclaving the media and cooling to 60°C.

β-Galactosidase assays were modified from published methods (22, 42, 49). Single colonies of yeast transformants from MinGlu plates were grown to saturation in 5 ml of MinGlu. Cells were then diluted into MinGlu (1:50), MinGluGal (1:50), MinGal (1:25), MinGalGlyEth (1:25), or MinGlyEth (1:10) and grown at 30°C until an optical density at 600 nm of about 1.0 was reached. Between 0.01 and 0.2 ml of culture was added to Z buffer (42) to give a total volume of 1.0 ml. Each sample then received 0.05 ml of 0.1% sodium dodecyl sulfate and 3 drops of chloroform and was votexed. The remainder of the assay and calculation of units of activity normalized for the optical density at 600 nm of the culture were exactly as described before (42). All assays were done in triplicate from independent transformants. Reproducibility was typically plus or minus 20% or less among triplicates. As little as 0.2 U was reproducibly detected in this assay.

RESULTS

DNA sequence. The sequence of the 914-base pair *Eco*RI fragment from pSc4816 containing the *GAL1-GAL10* divergent promoter region was determined (Fig. 1). All coordinates referred to hereafter in this paper will be the numbers assigned to base pairs in Fig. 1.

Coding regions. Translation of all six possible reading frames into amino acid sequences gave two open reading frames. These two divergent open reading frames begin near the ends of the sequenced fragment (base pairs 140 on the bottom strand and 821 on the top strand) and both run off

their respective ends. The open reading frame starting at base pair 821 matches 10 of the first 11 amino acids of a published protein sequence for galactokinase (52), which is encoded by *GAL1*. Threonine is the amino-terminal residue of the *GAL10* gene product epimerase (19). This is consistent with the ATG at base pair 140 being the actual *GAL10* translation start since the next codon is a threonine codon, ACA. Furthermore, the ATG at base pair 140 is the first ATG of the *GAL10* mRNA (see below), and translation usually begins at the first ATG of an mRNA in yeasts (57). Thus, the sequenced fragment contains 87 base pairs of *GAL1* coding sequence on the right and probably contains 140 base pairs of *GAL10* coding sequence on the left.

Base composition. Most of the 914 base pairs is adenine-thymine rich (65% overall), which is typical for *S. cerevisiae*. The genome of *S. cerevisiae* averages 60% adenine-thymine (15). However, bases 377 to 485 are rather guanine-cytosine rich (68%), which is unusual for this species.

Transcription initiation. Transcription start sites have been determined from in vivo mRNA by Johnston (27) for both *GAL1* and *GAL10* (see Fig. 1). Both genes have several starts. The major starts are at −13 for *GAL10* and −61 for *GAL1* with respect to the assumed translation start codons.

TATA boxes. The sequence TATAAA, or a closely related sequence, exists about 25 bases upstream from transcriptional starts in most higher eucaryotic genes that have been sequenced (D. Goldberg, Ph.D. thesis, Stanford University, Stanford, Calif., 1979). Such a sequence is also found upstream from yeast genes, but the distance from the start site is more variable and is usually farther than 25 bases (50, 55). The sequence TATAAG is found 114 bases upstream from the major *GAL10* transcription start and TATAAA is found 85 bases upstream from the major *GAL1* start. However, the adenine-thymine rich nature of these regions gives many other possibilities for these putative "TATA boxes." For example, TATTAA is found at −47 with respect to the *GAL10* start.

Comparison with *S. carlsbergensis*. The 914-base pair *GAL1-GAL10* control region from *S. cerevisiae* was compared with the equivalent region from *S. carlsbergensis* (6a). The two sequences were 95% homologous, with all of the differences occurring in the noncoding region (data not shown). Most of the differences were between bases 180 and 250 (see Fig. 1 for numbering), and there were no differences between bases 405 and 569.

Construction of *lacZ* fusions. Since the sequenced 914-base pair fragment overlapped both the *GAL1* and *GAL10* coding regions, precise, in-frame gene fusions of both *GAL1* and *GAL10* to *E. coli lacZ* could be constructed. This was accomplished by placing synthetic linkers, containing appropriate restriction sites, at both ends of the sequenced fragment. The fragment was then inserted in either orientation into available plasmids that contained the *lacZ* gene in a convenient form. These constructions resulted in genes coding for fused proteins containing the amino-terminal end of *GAL1* or *GAL10* and all but the extreme amino-terminal end of β-galactosidase. The construction of four pairs of these plasmids is shown in Fig. 2 to 5. Each pair contains one *GAL1* and one *GAL10* fusion. All eight plasmids are selectable in either *E. coli* or *S. cerevisiae*, but each pair has a different mode of replication or maintenance or both in *S. cerevisiae*. pRY131 and −133 contain a 2 μm circle origin of replication, whereas pRY161 and −163 contain a chromosomal origin of replication, *ARS1*, and the *CEN4* centromere. pRY171, −173, −181, and −183 contain no yeast origin of replication, and therefore they require integration into a chromosome for maintenance. pRY171 and −173 were directed to integrate at the *GAL1-GAL10* locus by cleaving with *Xho*I, wherease pRY181 and −183 were directed to integrate at the *LEU2* locus by cleaving with *Kpn*I.

All eight *lacZ* fusion plasmids formed uniformly blue colonies in *E. coli* P90C on 4′-bromo-5′-chlor-3′-indolyl-β-D-galactoside indicator plates, which is, in part, how they were screened during construction. However, only six of the eight *lacZ* fusion plasmids produced uniformly blue colonies in *S. cerevisiae* DBY745 on MinGal indicator plates (see Materials and Methods). pRY161 and −163 gave a range of colony color from pure white to dark blue under these conditions. Interestingly, pRY161 and −163 contain centromeres, which supposedly stabilize plasmid copy number (58). Hence, we could not satisfactorily explain this result. Due to the variable colony color on indicator plates conferred by pRY161 and pRY163, these plasmids were not used for any further studies.

Sequence of pRY121 fusion junction. pRY121 (and hence pRY131) contained an unexpected *Hin*dIII site near the *GAL1-lacZ* fusion junction. Therefore, the fusion junction was sequenced (Fig. 6). Evidently two *Hin*dIII linkers were inserted during the construction of pRY116 (see Fig. 2), and a single-base-pair deletion compensated in the subsequent construction of pRY121 to give an in-frame fusion. This artifact was not redressed, since the *Hin*dIII site proved to be a useful marker for the *GAL1* end of the insert (R. West, R. Yocum, and M. Ptashne, submitted for publication).

Regulation of *GAL-lacZ* fusions. To determine whether expression of the *GAL1*- and *GAL10-lacZ* fusions on the plasmids described above was regulated in the same manner as expression of the intact chromosomal *GAL1* and *GAL10* genes, the various plasmids were transformed into DBY745, which is wild type with respect to all *GAL* regulatory genes. β-Galactosidase activity was then determined after growth in repressing (MinGlu or MinGluGal), inducing (MinGal or MinGalGlyEth), or noninducing, nonrepressing (MinGlyEth) media. The results are shown in Table 2. All six plasmids showed the correct regulatory pattern: β-galactosidase was induced at least 1,000-fold by galactose, repressed at least 98-fold by glucose, and uninduced by the nonfermentable carbon sources glycerol plus ethanol.

To examine the effect of known *GAL* regulatory mutations on the expression of the *GAL-lacZ* fusions, a set of yeast strains was constructed to give four regulatory mutations, *gal4*, *GAL4*^{c}, *gal80*, and *GAL80*^{s}, in combination with *ura3* and *leu2* (see Table 1). This allowed selection of the fusion plasmids in the regulatory mutant strains. Each strain was transformed with the *GAL1-lacZ* fusion plasmid pRY131, grown in the five media mentioned above, and assayed for β-galactosidase activity (Table 3). As expected, the *gal4* and the *GAL80*^{s} mutants gave little or no activity in any medium, whereas both *GAL4*^{c} and *gal80* led to constitutive expression of the fusion in nonrepressing, noninducing medium and partially constitutive expression in repressing media.

In vitro deletions. A 365-base pair *Dde*I-to-*Sau*3AI fragment (bases 298 to 663 in Fig. 1) was previously shown to contain sufficient information for galactose regulation (the *GAL* UAS) by inserting this fragment into the 5′ flanking region of a *CYC1-lacZ* fusion (23). We further delimited the sequences that were sufficient to enable galactose induction to the *GAL10* side of the UAS by constructing a set of deletions in vitro (see Materials and Methods). We chose pLGSD5 (23) as the parent plasmid for constructing the

```
        10         20         30         40         50
GAATTCGACA GGTTATCAGC AACAACACAG TCATATCCAT TCTCAATTAG
CTTAAGCTGT CCAATAGTCG TTGTTGTGTC AGTATAGGTA AGAGTTAATC

        60         70         80         90        100
CTCTACCACA GTGTGTGAAC CAATGTATCC AGCACCACCT GTAACCAAAA
GAGATGGTGT CACACACTTG GTTACATAGG TCGTGGTGGA CATTGGTTTT

       110        120        130        140        150
CAATTTTAGA AGTACTTTCA CTTTGTAACT GAGCTGTCAT TTATATTGAA
GTTAAAATCT TCATGAAAGT GAAACATTGA CTCGACAGTA AATATAACTT
                                  GAL 10
       160        170                   190        200
TTTTCAAAAA TTCTTACTTT TTTTTTGGAT GGACGCAAAG AAGTTTAATA
AAAAGTTTTT AAGAATGAAA AAAAAACCTA CCTGCGTTTC TTCAAATTAT
    *          ***                     *          *

       210                                         250
ATCATATTAC ATGGCATTAC CACCATATAC ATATCCATAT ACATATCCAT
TAGTATAATG TACCGTAATG GTGGTATATG TATAGGTATA TGTATAGGTA

       260        270        280        290        300
ATCTAATCTT ACTTATATGT TGTGGAAATG TAAAGAGCCC CATTATCTTA
TAGATTAGAA TGAATATACA ACACCTTTAC ATTTCTCGGG GTAATAGAAT

       310        320        330        340        350
GCCTAAAAAA ACCTTCTCTT TGGAACTTTC AGTAATACGC TTAACTGCTC
CGGATTTTTT TGGAAGAGAA ACCTTGAAAG TCATTATGCG AATTGACGAG

       360        370        380        390        400
ATTGCTATAT TGAAGTACGG ATTAGAAGCC GCCGAGCGGG TGACAGCCCT
TAACGATATA ACTTCATGCC TAATCTTCGG CGGCTCGCCC ACTGTCGGGA

                  420        430        440        450
CCGAAGGAAG ACTCTCCTCC GTGCGTCCTC GTCTTCACCG GTCGCGTTCC
GGCTTCCTTC TGAGAGGAGG CACGCAGGAG CAGAAGTGGC CAGCGCAAGG

       460        470        480        490        500
TGAAACGCAG ATGTGCCTCG CGCCGCACTG CTCCGAACAA TAAAGATTCT
ACTTTGCGTC TACACGGAGC GCGGCGTGAC GAGGCTTGTT ATTTCTAAGA

       510        520        530        540        550
ACAATACTAG CTTTTATGGT TATGAAGAGG AAAAATTGGC AGTAACCTGG
TGTTATGATC GAAAATACCA ATACTTCTCC TTTTTAACCG TCATTGGACC

       560        570        580        590        600
CCCCACAAAC CTTCAAATGA ACGAATCAAA TTAACAACCA TAGGATGATA
GGGGTGTTTG GAAGTTTACT TGCTTAGTTT AATTGTTGGT ATCCTACTAT

       610        620        630        640        650
ATGCGATTAG TTTTTTAGCC TTATTTCTGG GGTAATTAAT CAGCGAAGCG
TACGCTAATC AAAAAATCGG AATAAAGACC CCATTAATTA GTCGCTTCGC

       660        670        680        690        700
ATGATTTTTG ATCTATTAAC AGATATATAA ATGCAAAAAC TGCATAACCA
TACTAAAAAC TAGATAATTG TCTATATATT TACGTTTTTG ACGTATTGGT

       710        720        730        740        750
CTTTAACTAA TACTTTCAAC ATTTTCGGTT TGTATTACTT CTTATTCAAA
GAAATTGATT ATGAAAGTTG TAAAAGCCAA ACATAATGAA GAATAAGTTT

  **      *   **                 780        790        800
          *
TGTAATAAAA GTATCAACAA AAAATTGTTA ATATACCTCT ATACTTTAAC
ACATTATTTT CATAGTTGTT TTTTAACAAT TATATGGAGA TATGAAATTG

       810        820    GAL 1              840        850
GTCAAGGAGA AAAAACTATA ATGACTAAAT CTCATTCAGA AGAAGTGATT
CAGTTCCTCT TTTTTGATAT TACTGATTTA GAGTAAGTCT TCTTCACTAA

       860        870        880        890        900
GTACCTGAGT TCAATTCTAG CGCAAAGGAA TTACCAAGAC CATTGGCCGA
CATGGACTCA AGTTAAGATC GCGTTTCCTT AATGGTTCTG GTAACCGGCT

       910
AAAGTGCGGA ATTC   Eco RI  Linker
TTTCACGCCT TAAG
```

FIG. 1. DNA sequence of the *GAL1-GAL10* divergent promoter region. The 914-base pair *Eco*RI fragment from pSc4816 was the source of DNA. The *Eco*RI site at base 908 was formerly an *Ava*I site in the amino-terminal end of the *GAL1* gene. Probable translation start codons are indicated by arrowhead boxes. Possible "TATA" sequences are enclosed in plain boxes. Selected short repeats are shown with arrows. Transcription start sites are indicated by * (27).

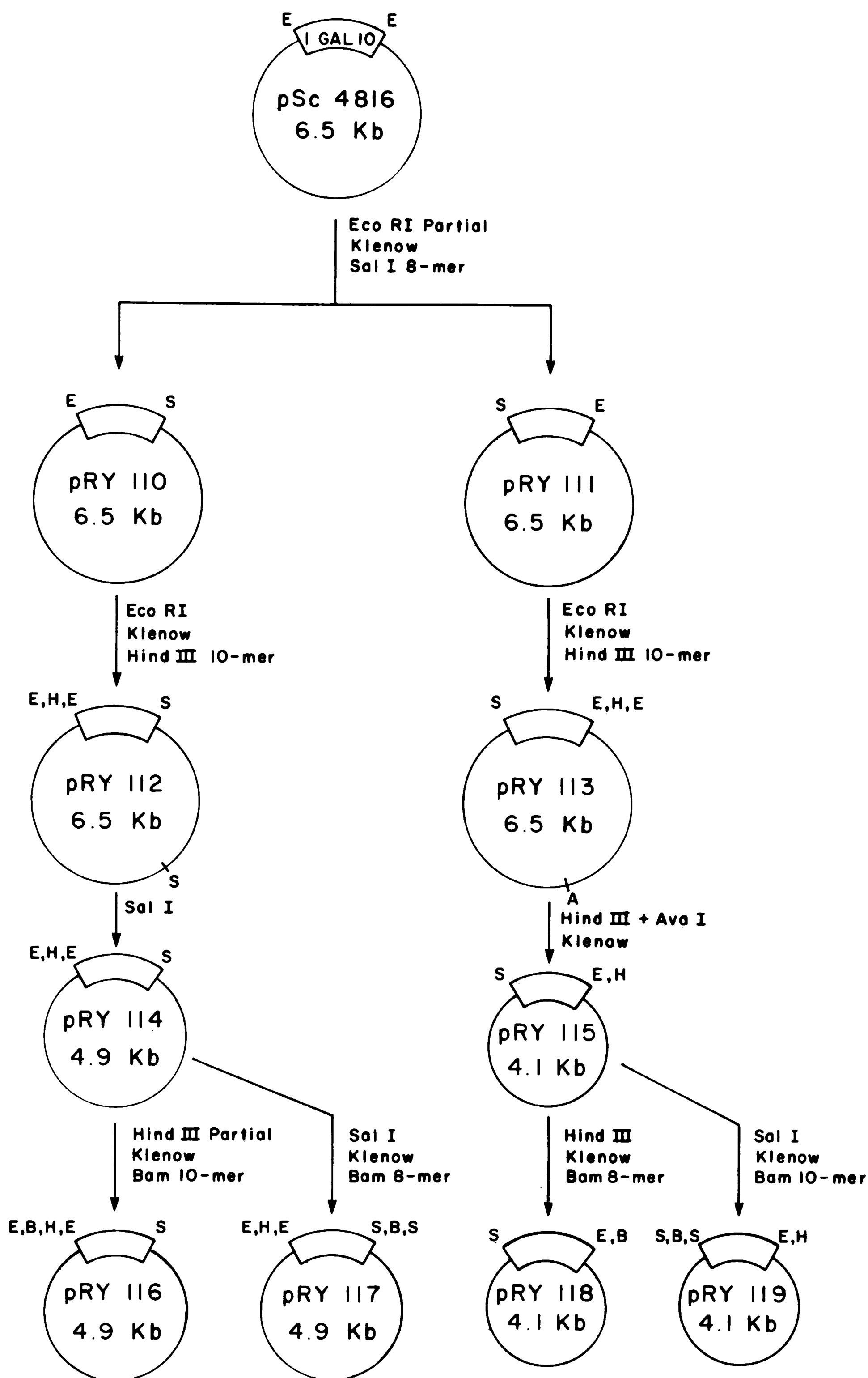


FIG. 2. Construction of intermediate plasmids. pRY116, -117, -118, and -119 were designed to provide truncated *GAL1* and *GAL10* genes and the intergenic sequence for in-frame fusions with *E. coli lacZ* in pLG669 and pSZ211. Abbreviations: A, *Ava*I; B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; S, *Sal*I.

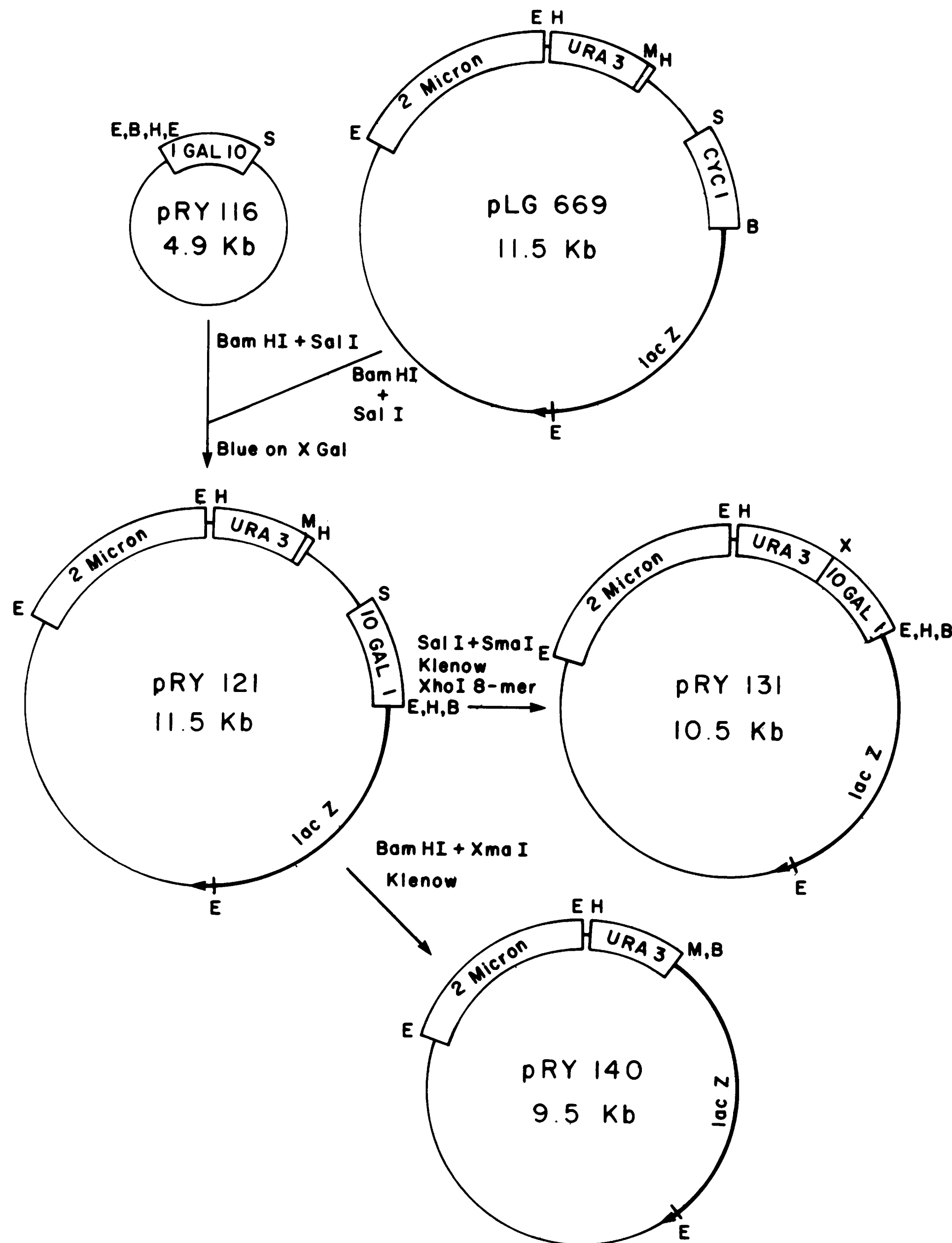


FIG. 3. Construction of *GAL-lacZ* fusions on plasmids containing the 2 μm replication origin. Shown is construction of pRY131. Construction of pRY133 was identical except that the starting plasmid was pRY118 instead of pRY116 (see Fig. 2). pRY140 was designed to receive internal deletions of the divergent promoter constructed in vitro (West et al., submitted for publication). Abbreviations: B, *Bam*HI; E, *Eco*RI, H, *Hin*dIII; M, *Xma*I or *Sma*I; S, *Sal*I; X, *Xho*I; Kb, kilobases.

deletions since it contained a unique *Xho*I site at the border of the 365-base pair *GAL* UAS. The deletions progressively removed the *GAL10* side of the 365-base pair insert starting at base pair 298 (Fig. 7). The deleted plasmids were then transformed into DBY745, and the transformants were assayed for β-galactosidase activity (see Fig. 7). A deletion from base 298 to base 396 or 406 (Δ18 or Δ16) showed no significant loss of activity upon galactose induction. A deletion to base 452 (Δ31) showed greatly reduced, but nonetheless measurable, induction, whereas a deletion to base 507 (Δ34) showed complete loss of inducible fusion activity.

Homologous repeats. To locate possible regulatory sites in the *GAL1-GAL10* control region, a search for regions of dyad symmetry, internal homologies, and homologies with the region just 5′ to the *GAL7* gene (44) was made with the MOLGEN<SEQ> computer program. The homology parameters were set to the least possible stringency: homology minimum at 50%, minimum number of consecutive matches at three bases, and probability of significance at 95%. Three types of statistically significant repeated sequences were found: direct repeats, inverted repeats, and short sequences found in both the *GAL1-GAL10* and the *GAL7* promoter regions. A summary of these repeated sequences showing

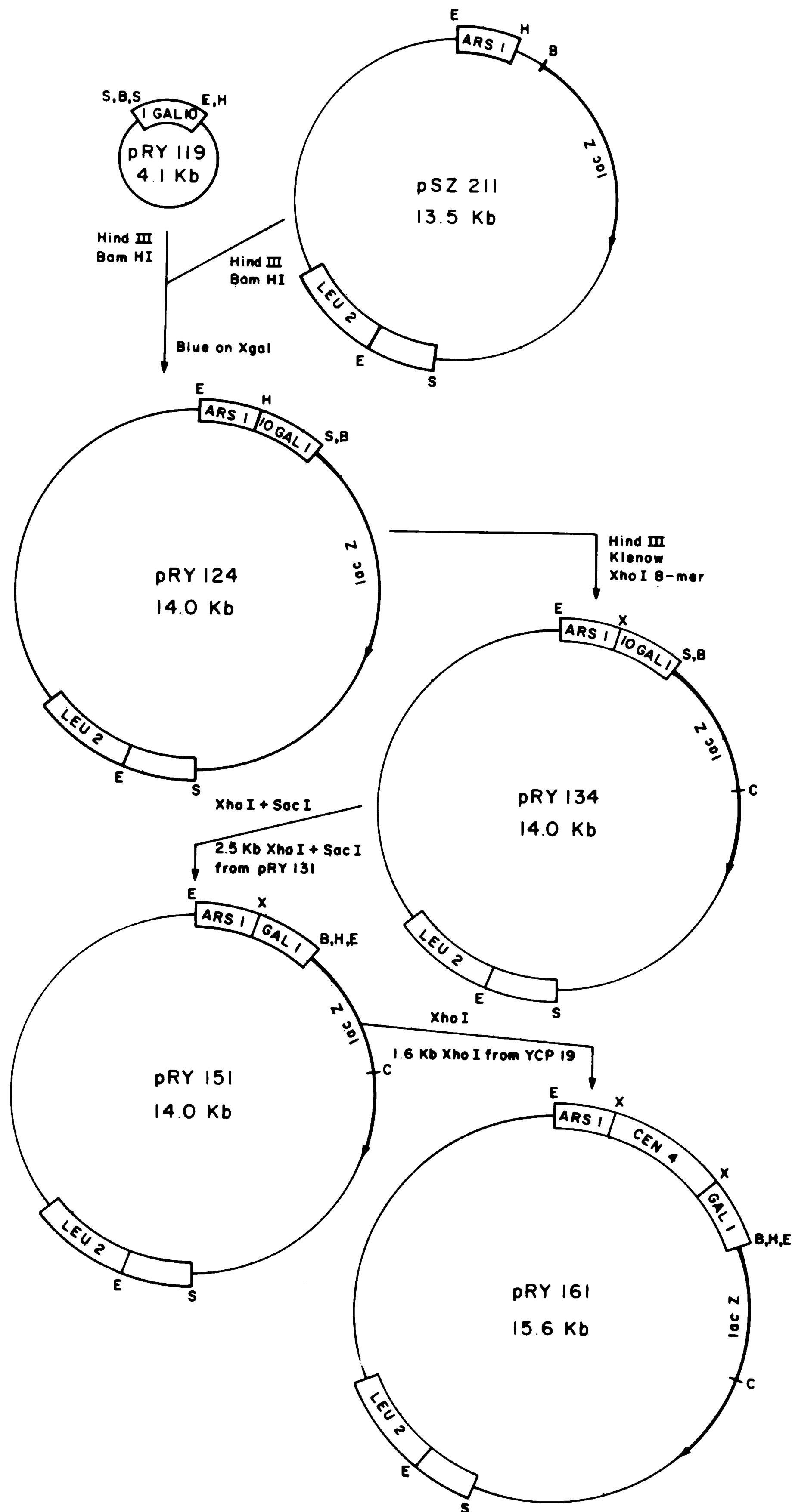


FIG. 4. Construction of *GAL-lacZ* fusions on plasmids containing the chromosomal origin of replication, *ARS1*, and the centromere, *CEN4*. Shown is construction of pRY161. Construction of pRY163 was identical except that the starting plasmid was pRY117 instead of pRY119 (see Fig. 2). Abbreviations: B, *Bam*HI; C, *Sac*I; E, *Eco*RI; H, *Hin*dIII; S, *Sal*I; X, *Xho*I; Kb, kilobases.

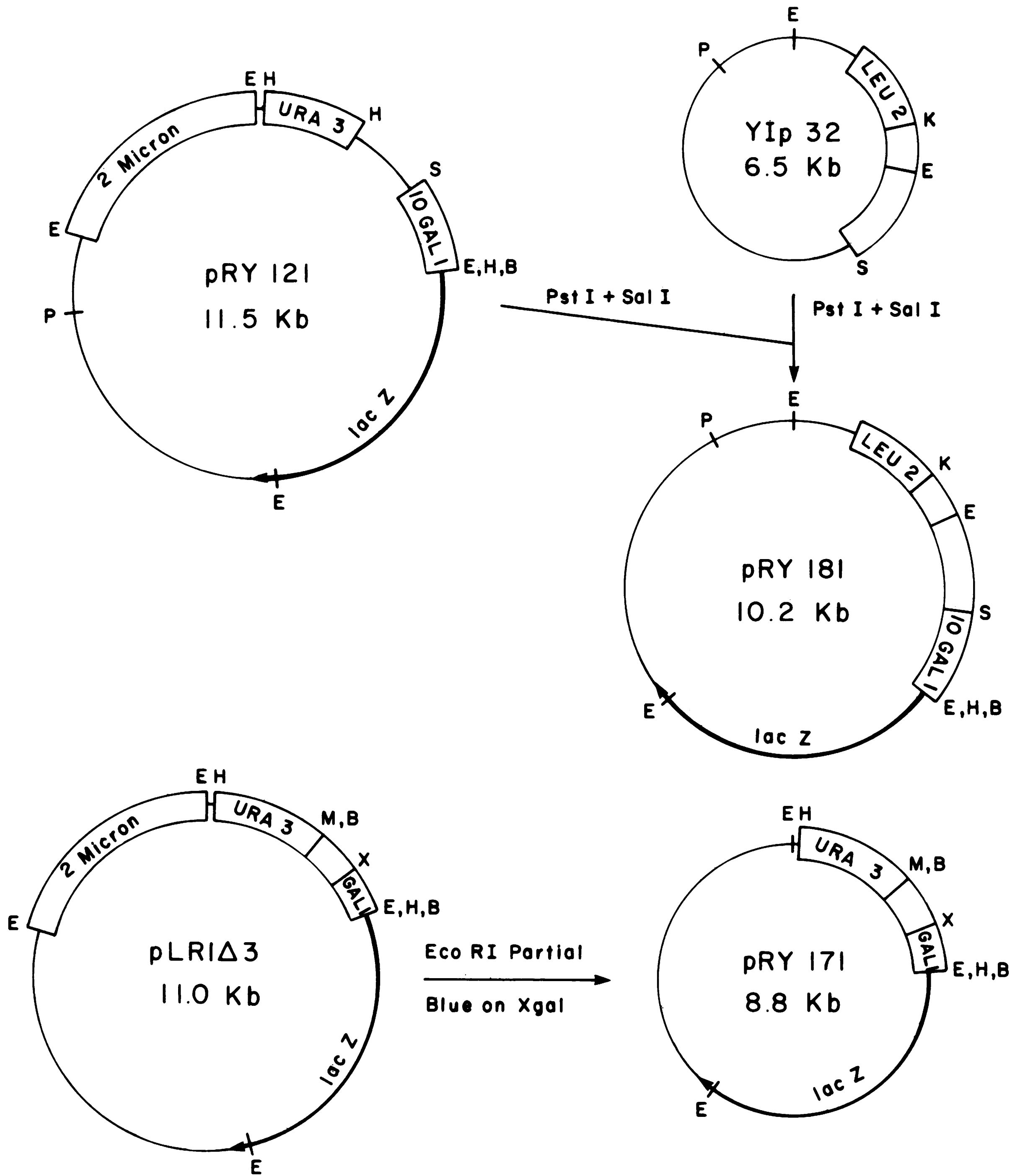


FIG. 5. Construction of *GAL-lacZ* fusions on integrating plasmids. Starting plasmid pLR1Δ3 is similar to pRY131 (Fig. 3) except that a *Xho*I linker has been artificially inserted in the middle of the divergent promoter (West et al., submitted for publication). Shown is construction of pRY181 and pRY171. Construction of pRY183 was identical to that of pRY181 except that the starting plasmid was pRY123 (see Fig. 3). pRY173 was constructed by reversing the orientation of the *Bam*HI fragment of pRY171 so that an in-frame *GAL10-lacZ* fusion resulted. Abbreviations: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; M, *Xma*I; P, *Pst*I; S, *Sal*I; X, *Xho*I; Kb, kilobases.

their positions upstream from *GAL1*, *GAL7*, and *GAL10* (lined up at their respective major transcription start sites) is shown in Fig. 8. Only a subset of these elements occurred at similar locations in the various genes. Sequences A, B, C, C', E, F, and G varied substantially in distance from the start sites of the genes in which they were found. The highly significant 12-base pair repeat, G, was in the region that matched most poorly with *S. carlesbergensis*, and in fact neither repeat was intact in *S. carlsbergensis* (data not shown). Furthermore, all of these sequences from *GAL1* and *GAL10* lie outside of sequences that are known to be sufficient to confer galactose induction and glucose repres-

sion upon an adjancent promoter (see above and reference 23). Sequence I and J are almost identically situated with respect to the *GAL1* and *GAL10* starts. However, these sequences are also outside the region that is sufficient to confer galactose induction toward the *GAL10* side (see above), and neither site exists upstream from *GAL7*. Thus, it is difficult to imagine how any of these sequences could be involved in galactose regulation.

The direct repeat D and the inverted repeat H are roughly the same distance from transcription start sites of *GAL1* and *GAL10* and lie within the region suspected to contain the regulatory sequences (see above). A single copy of site D is found upstream from *GAL7*. Perhaps these sites are involved in galactose regulation. They might represent binding sites for *GAL4* protein or other regulatory proteins such as the positive regulator encoded by *GAL11* (43).

DISCUSSION

Characterization of *GAL-lacZ* fusions. The fusion of a slightly truncated *E. coli lacZ* gene to the amino-terminal coding sequences of a yeast gene gives an easily assayed gene product whose expression is regulated via the flanking yeast sequences (22, 49). Such fusions, constructed on plasmids, greatly facilitate the study of effects of mutations introduced into regulatory sequences. Before undertaking such a study of the regulation of *GAL1* and *GAL10*, we wanted to be certain that regulation of *GAL1*- and *GAL10-lacZ* fusions on plasmids accurately mimicked regulation of intact, chromosomal *GAL1* and *GAL10*. Regulation of both *GAL1*- and *GAL10-lacZ* fusions on integrating plasmids and plasmids bearing the 2 μm replicator accurately reflected the regulation pattern of native *GAL1* and *GAL10*. All fusions showed >1,000-fold induction by galactose and between 90- and 180-fold repression by glucose plus galactose (see Table 2). For comparison, galactokinase (*GAL1*) is induced 1,000-fold by galactose and repressed 1,000-fold by glucose (37). The difference between the extent of glucose repression of our fusions and chromosomal *GAL1* might result from differences between the media used in the two studies. We routinely used minimal selective media containing 2% galactose plus or minus 2% glucose to grow plasmid-containing yeasts, whereas the studies on galactokinase used a rich medium containing 1% galactose plus or minus 5% glucose (37). The increased ratio of glucose to galactose could account for the increased glucose repression found in the latter study.

For all pairs of matched *GAL1* and *GAL10* fusion plasmids, the *GAL1* fusion gave two- to fourfold more activity than the corresponding *GAL10* fusion (see Table 2). This is consistent with galactokinase being synthesized at four times the level of epimerase from the native chromosomal genes (19, 53).

Comparison of β-galactosidase activity from different types of vectors leads to several important conclusions. Integrated fusions gave the same pattern of regulation as fusions borne on replicating circular plasmids, ruling out the possibility that the observed galactose induction was a result of increased plasmid copy number (see Table 2). The nearly

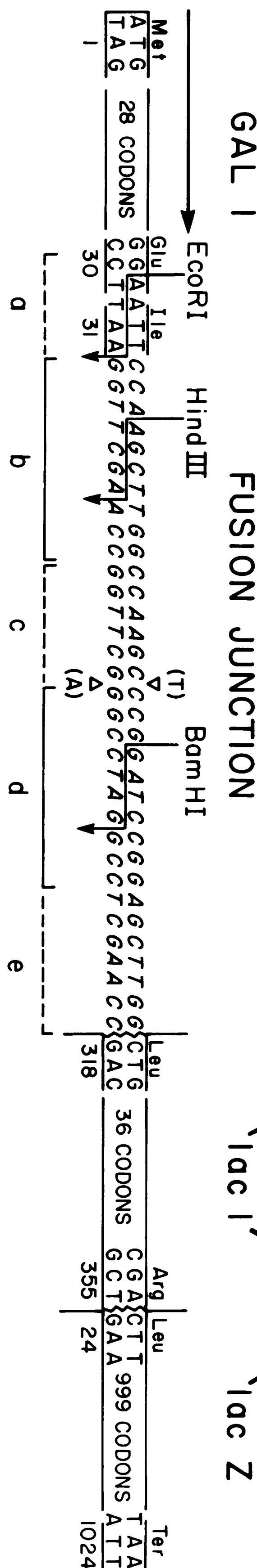


FIG. 6. Sequence of the fusion junction of pRY121 and pRY131. The letters a to e indicate synthetic linkers or remnants thereof that were introduced at the fusion junction. 'lacI' refers to a short portion of the *lacI* gene that occurs in all plasmids derived from pLG669 (20, 22).

TABLE 2. β-Galactosidase activity of *GAL1*- and *GAL10-lacZ* fusion plasmids in *S. cerevisiae* DBY745 grown in minimal selective media containing various carbon sources

Plasmid	Fusion type	Plasmid type	β-Galactosidase activity (U)					Glucose repression ratio[a]
			MinGlu	MinGlu Gal	MinGal	MinGly Eth	MinGal GlyEth	
pRY131	*GAL1*	2μm	0	83	8,110	0	8,190	98
pRY133	*GAL10*	2μm	0	10	1,810	0	1,860	181
pRY171	*GAL1*	Integrated at *GAL1*	0	29	3,800	0	3,650	131
pRY173	*GAL10*	Integrated at *GAL10*	0	12	1,310	0	1,260	108
pRY181	*GAL1*	Integrated at *LEU2*	0	25	3,120	0	3,180	125
pRY183	*GAL10*	Integrated at *LEU2*	0	9.0	1,220	0	1,230	135

[a] Glucose repression ratio is defined as the number of units from cells grown in galactose divided by the number of units from cells grown in glucose plus galactose.

identical behavior of fusions integrated at the *GAL1-GAL10* locus (pRY171 and -173) and fusions integrated at the *LEU2* locus (pRY181 and -183) indicated that all sequences necessary for proper regulation are contained on the 914-base pair *GAL* segment and that long-range effects are probably unimportant. If a change in chromatin structure accompanies galactose induction, then all information needed for that change exists on the 914-base pair fragment.

Two recent determinations of the copy number in yeasts of plasmids containing the 2 μm replicator gave 7 and 14 copies per haploid genome (24, 28). If we assume that pRY131 exists at seven copies per haploid genome, then the relationship between gene copy number and β-galactosidase activity is nonlinear. pRY171 integrated at one copy per genome gives about half the activity of multipcopy pRY131 (see Table 2). The *GAL4* protein is probably present at only a few copies per cell (30, 31), so it is likely that *GAL4* is limiting and is titrated out by a multicopy plasmid.

Mutations in *GAL4* and *GAL80* altered expression of the *GAL1-lacZ* fusion on pRY131 as predicted from the responses of chromosomal *GAL1* and *GAL10* to the same mutations (11, 47). However, one surprising result was that full glucose repression occurred in our *gal80* mutant (Table 3). This suggests that the *GAL80* protein plays no role in glucose repression. This contradicts two previous studies, which concluded that the *GAL80* protein mediates glucose repression (28) and that inducer exclusion partially accounts for glucose repression in a *GAL80* strain, but not in a *gal80* strain (23). Resolution of these disagreements will require a strain deleted for *GAL80*, since the exact allele of *gal80* used in the present study was unknown, and it is possible that this allele was only partially deficient for *GAL80* function. Glucose repression of the *GAL1-lacZ* fusion on pRY131 was also seen in our *GAL4*[c] strain. Thus, constitutive mutations in both *GAL80* and *GAL4* can still allow normal glucose repression. One interpretation of these results is that glucose repression is, to some extent, independent of the *GAL4-GAL80* induction system. However it is also possible that *GAL4* or *GAL80* or both have functionally independent domains that carry out glucose repression and galactose induction and that our mutant alleles affect only the latter.

We have no good explanation for the growth of our *GAL80*[s]/pRY131 strain in MinGal medium. Perhaps the *GAL80*[s]*-96* allele used in this study is leaky as was found for a different allele, *GAL80*[s]*-16* (12).

The hyperinduction of *GAL1* from pRY131 in our *GAL4*[c] strain grown in glycerol plus ethanol (18,000 versus 8,000 U in galactose; see Table 3) might result from galactose being a partially repressing carbon source in *S. cerevisiae*. In *E. coli*, sugars closely related to glucose, such as galactose and lactose, are known to exhibit partial catabolite repression (33). Hyperinduction of galactokinase in a *GAL4*[c] strain has been previously reported (45). Hyperinduction of *GAL1* from pRY131 was also found in our *gal80* strain grown in galactose (19,000 versus 10,000 U in glycerol plus ethanol; see Table 3). This might result from partial *GAL80* function in our *gal80* strain or *GAL80*-independent activation of *GAL4* by galactose or both.

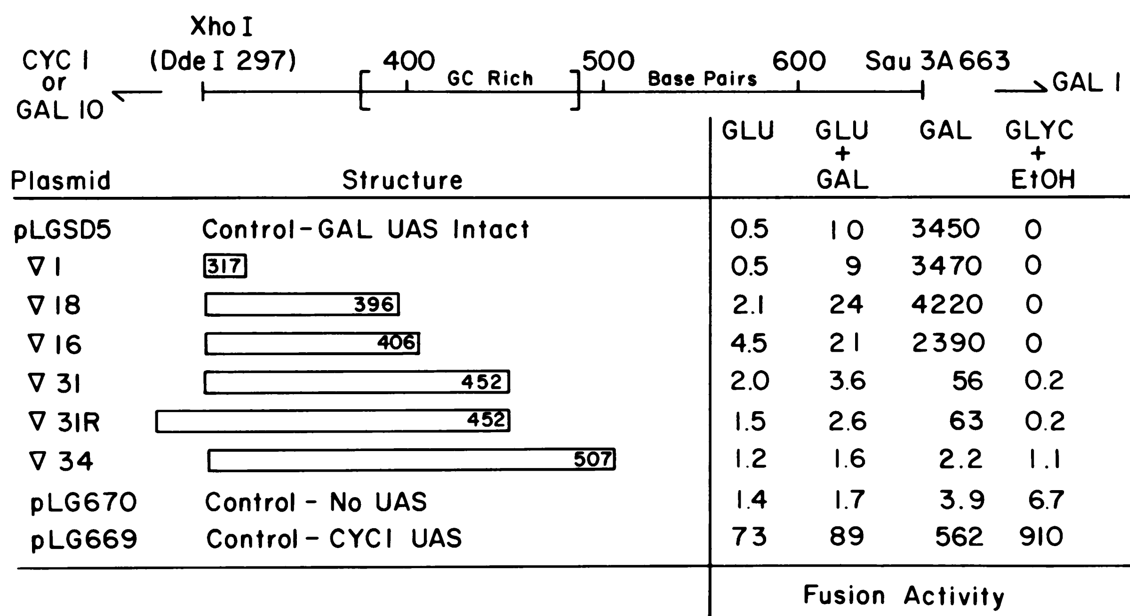
In vitro deletions. The regulatory sequences in the *GAL1-GAL10* intergenic region that respond to *GAL4* had been previously shown to reside in a 365-base pair fragment that lies about midway between *GAL1* and *GAL10* (23). The series of deletions shown in Fig. 7 further defines the minimal sequences sufficient for galactose induction to the *GAL10* side (see Results). We conclude from these deletions that there are two adjacent regions involved in galactose regulation: (i) base pairs 406 to 451, which contain sequences necessary for full induction; and (ii) base pairs 452 to 506, which contain sequences capable of independently conferring a low level of galactose induction. The role of the latter sequences in regulation by an intact *GAL* UAS could be minor. In fact, these results do not necessarily imply that there are two sites of *GAL4* action. The region from base pairs 452 to 506 could simply contain part of a single site.

We doubt that the distance between the *GAL* UAS sequences and transcription start sites is critical, since Δ16 and Δ18 bring the UAS about 100 base pairs closer to the transcriptions starts without significantly affecting expression of β-galactosidase. Furthermore, Δ31 and Δ31R show nearly identical levels of β-galactosidase, even though Δ31R brings the truncated UAS 23 base pairs closer to the transcription starts than does Δ31. Thus, we attribute all differ-

TABLE 3. β-Galactosidase activity of pRY131 in several *GAL* regulatory mutants grown in minimal selective media containing various carbon sources

Strain	*GAL* genotype	β-Galactosidase activity (U)				
		MinGlu	MinGlu Gal	MinGal	MinGly Eth	MinGal GlyEth
DBY745	*GAL4*[+] *GAL80*[+]	0	110	7,370	0	7,940
SHC22c	*gal4 GAL80*[+]	0	0	—[a]	0	0
SHI6a	*GAL4*[+] *GAL80*[s]	0	0	7[b]	3	17
SHE2b	*GAL4*[+] *gal80*	37	59	18,900	10,400	19,800
SHJ1d	*GAL4*[c] *GAL80*[+]	10	155	8,370	18,000	9,420

[a] —, Did not grow at all.
[b] Grew very slowly.



Plasmid	Structure	GLU	GLU + GAL	GAL	GLYC + EtOH
pLGSD5	Control-GAL UAS Intact	0.5	10	3450	0
∇ 1	317	0.5	9	3470	0
∇ 18	396	2.1	24	4220	0
∇ 16	406	4.5	21	2390	0
∇ 31	452	2.0	3.6	56	0.2
∇ 31R	452	1.5	2.6	63	0.2
∇ 34	507	1.2	1.6	2.2	1.1
pLG670	Control - No UAS	1.4	1.7	3.9	6.7
pLG669	Control - CYCI UAS	73	89	562	910

FIG. 7. Deletions of the *GAL* UAS and their effect on activation of transcription to the *GAL10* side. Deletions of pLGSD5 (23) from the *Xho*I site rightward into the *GAL* sequences were constructed as described in the text. Δ31R contained an accidental additional short deletion to the left. As controls, pLG669 and pLG670, which contain the *CYC1* UAS and no UAS, respectively, were also included (22). The scale at the top refers to base pair coordinates of Fig. 1. The rectangular boxes indicate the sequences deleted from the bar shown on top. The numbers in the rectangular boxes give the coordinate of the first base pair of the deletion endpoint. The brackets on the bar show the guanine-cytosine (GC)-rich region. The numbers to the right give units of β-galactosidase activity for each deleted plasmid in DBY745 grown in minimal selective media containing the carbon sources indicated (see text).

ences between these particular deletions to presence or absence of sequences rather than to spacing between elements.

All of the deletions shown in Fig. 7 gave parallel responses of galactose induction and glucose repression, regardless of the extent of deletion or extent of galactose induction. Even Δ31, which is severely impaired for galactose induction, still showed 18-fold glucose repression. Thus, we were unable to separate the two types of regulation. Either both types of regulation utilize the same DNA site, or else the sequences that govern glucose repression to the *GAL10* side lie further away from *GAL10* than the sequences involved in galactose activation.

The sequences shown by our deletions to be involved in galactose induction and glucose repression (base pairs 406 to 507) are a subset of the sequences totally conserved between *S. cerevisiae* and *S. carlsbergensis* (base pairs 406 to 568). This is consistent with the involvement of these sequences in regulation. The two species are closely enough related so that the *S. cerevisiae* regulatory proteins operate correctly on the *S. carlsbergensis GAL1,7,10* genes (6a). These implicated regulatory sequences substantially overlap the guanine-cytosine-rich region noted in Results (base pairs 377 to 485). We know of no good reason why regulatory sequences should be enriched for guanine-cytosine.

pRY121 and -123 were designed so that a set of 5′ deletions of both the *GAL1* and *GAL10* fusions could be easily constructed. They were also designed for easily combining these deletions in vitro in pRY140 (see Fig. 3) to give a set of internal deletions and "linker" mutations (41). An extensive set of such mutations has been constructed and analyzed and will be presented elsewhere (West et al., submitted for publication).

Repeated DNA sequences. There are several well-known examples of positive and negative control of gene transcription by regulatory proteins that bind to DNA sites at or near the initiation point for transcription. Some of these regulatory regions contain divergent promoters. For example, the right operators of bacteriophages lambda, P22, and 434 all contain three tandem rotationally symmetric sites that bind repressor or cro to regulate divergent transcription from p_R and p_{RM} (26). Transcription from the early and late promoters of simian virus 40 is regulated by binding of T antigen to three tandem sites located between the early and late transcription starts (8). In addition, adjacent to the three T-antigen binding sites are a pair of directly repeated 21-base pair sites that bind a regulatory protein called Sp1 (14) and a 72-base pair direct repeat that contains a transcription "enhancer" element that is necessary for high levels of transcription from the early promoter (18). All of these regulatory sites are recognizable by features of their primary sequence structures. We therefore examined the *GAL1-GAL10* divergent promoter region and the *GAL7* promoter region (44) for common sequences.

Of the many short homologies found, only a couple, D and H, were in the region between bases 406 and 507 that was shown by our deletions to be involved in galactose and glucose regulation (see Results and Fig. 7 and 8). It is tempting to speculate that the inverted repeat H, GAAGACN$_{18}$GTCTTC, might be a binding site for the *GAL4* regulatory protein. The site is rotationally symmetric, as are many known regulatory protein binding sites (26). Furthermore, its 30-base pair size should be large enough to accommodate the *GAL4* protein, which is predicted from its DNA sequence to be 99,000 daltons (31a). Simian virus 40 T antigen, which is 94,000 daltons, binds to 30-base pair sites (8). The inverted repeat H is not found in the *GAL7* upstream sequence, but a sequence that resembles it, GAAGGCN$_{18}$GTCTAA, was found 250 base pairs upstream from the *GAL7* transcription start. This spacing is nearly

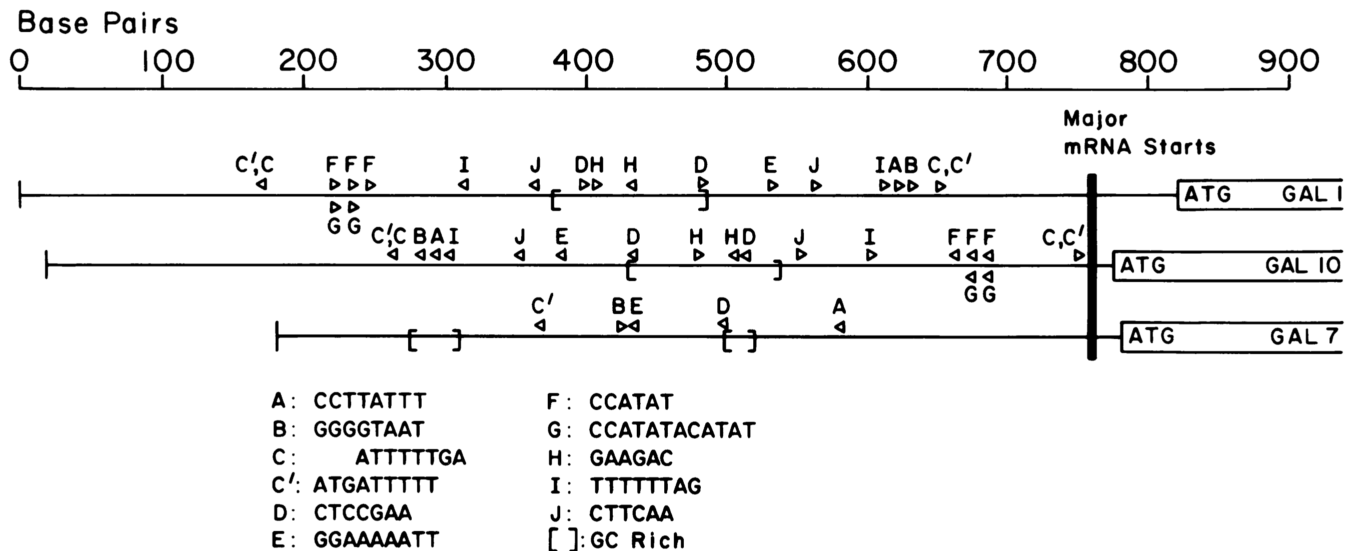


FIG. 8. Schematic diagram of short DNA sequence homologies in the upstream regions of *GAL1,7,10*. The three upstream DNA sequences were compared pairwise with the MOLGEN<SEQ> homologies program. The sequences were lined up at their respective major transcription start sites. The scale at the top indicates the base pair coordinates from Fig. 1 lined up with the *GAL1* gene. The *GAL7* sequences and major transcription start site were provided by Nogi and Fukasawa (44). The position of each sequence is indicated by an arrowhead pointing in the 5' to 3' direction of the sequence referred to by the capital letter. A list of the actual sequences is given below the schematic. GC, guanine-cytosine.

identical to that of the H repeats and the *GAL10* starts (see Fig. 8).

Of course, it is possible that the homology search did not uncover the actual *GAL4* binding sites. As a control experiment, the bacteriophage lambda and 434 right and left operators were subjected to the same homology search program used above. None of the operator sites were picked out (data not shown). This is because (i) the lambda sites have imperfect symmetry within the sites and spotty homology between sites, and (ii) the conserved 4-base-pair motif at the ends of the phage 434 operator sites was not considered statistically significant by the program used. Thus, well-known symmetric protein binding sites are not necessarily found by computer searches. The role, if any, for the repeated sequences found in the computer search will require further study.

Conclusion. We have constructed *GAL1*- and *GAL10-lacZ* fusions and have shown that these fusions can be used to study the regulation of *GAL1* and *GAL10*. We used these fusions to identify a 45-base pair sequence midway between *GAL1* and *GAL10* that contains information necessary for a normal level of galactose induction toward the *GAL10* side. This sequence lies between -253 and -298 with respect to the major *GAL10* transcription starts. From -299 to -353 is another region containing sequences that can independently confer a measurable but low level of galactose induction to the *GAL10* side. Our results are consistent with those found in a similar study of the effect of deletions on the expression of chromosomal *GAL1* and *GAL10* (27). Future work will address the mechanisms whereby these regulatory sequences control transcription from sites hundreds of base pairs away from transcription starts.

ACKNOWLEDGMENTS

We thank Mark Johnston for giving us data prior to publication, especially for the start points of *GAL1* and *GAL10* transcription, Thomas St. John for pSc4816, Leonard Guarente for pLG699 and pLG670, Jack Szostak for pSZ211, Carl Mann for YCp19, Bruce Citron for *S. carlsbergensis GAL* sequences, Yasuhisa Nogi and Toshio Fukasawa for the *S. cerevisiae GAL7* upstream sequence and transcription start site, James Yarger for several essential yeast strains, Allan Laughon for the *GAL4* gene sequence, Stephen Johnston for help in tracing regulatory allele numbers, and Pam Silver for critically reading the manuscript.

This work was supported in part by Public Health Service grants AI18589 and GM32308 to Mark Ptashne from the National Institutes of Health and in part by Biotechnica International, Cambridge, Mass. During part of this work R.R.Y. was a Helen Hay Whitney Postdoctoral Fellow.

LITERATURE CITED

1. **Adams, B.** 1972. Induction of galactokinase in *Saccharomyces cerevisiae*: kinetics of induction and glucose effects. J. Bacteriol. **111**:308–313.
2. **Beier, D., and E. Young.** 1982. Characterization of a regulatory region upstream of the ADR2 locus of *S. cerevisiae*. Nature (London) **300**:724–728.
3. **Botstein, D., S. Falco, S. Stewart, M. Brennan, S. Scherer, D. Stinchcomb, K. Struhl, and R. Davis.** 1979. Sterile host yeast (SHY): a eukaryotic system of biological containment for recombinant DNA experiments. Gene **8**:17–24.
4. **Broach, J.** 1979. Galactose regulation in *Saccharomyces cerevisiae*: the enzymes encoded by the *GAL7,10,1* cluster coordinately controlled and separately translated. J. Mol. Biol. **131**:41–56.
5. **Carlson, M., B. Osmond, and D. Botstein.** 1981. Mutants of yeast defective in sucrose utilization. Genetics **98**:25–46.
6. **Cirillo, P.** 1968. Galactose transport in *Saccharomyces cerevisiae*: nonmetabolizable sugars as substrates and inducers of the galactose transport system. J. Bacteriol. **95**:1727–1731.

6a. **Citron, B., and J. Donelson.** 1984. Sequence of the *Saccharomyces GAL* region and its transcription in vivo. J. Bacteriol. **158**:269–278.

7. **Citron, B., M. Feiss, and J. Donelson.** 1979. Expression of the yeast galactose gene in *Escherichia coli*. Gene **6**:251–264.
8. **DiMiao, D., and D. Nathans.** 1982. Regulatory mutants of simian virus 40: effects of mutations at a regulatory site on DNA replication and expression of viral genes. J. Mol. Biol.

156:531–548.
9. **Donahue, T., R. Daves, G. Lucchini, and G. Fink.** 1983. A short nucleotide sequence required for regulation of HIS4 by the general control system of yeast. Cell **32:**89–98.
10. **Douglas, H., and D. Hawthorne.** 1964. Enzymatic expression and genetic linkage of genes controlling galactose utilization in *Saccharomyces*. Genetics **49:**837–844.
11. **Douglas, H., and D. Hawthorne.** 1966. Regulation of genes controlling synthesis of the galactose pathway enzymes in yeast. Genetics **54:**911–916.
12. **Douglas, H., and D. Hawthorne.** 1972. Uninducible mutants in the *gal i* locus of *Saccharomyces cerevisiae*. J. Bacteriol. **109:**1139–1143.
13. **Douglas, H., and G. Pelroy.** 1963. A gene controlling inducibility of the galactose pathway enzymes in *Saccharomyces*. Biochim. Biophys. Acta **68:**155–156.
14. **Dynan, M., and R. Tjian.** 1983. The promoter specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell **35:**79–87.
15. **Fangman, W., and V. Zakian.** 1982. Genome structure and replication, p. 27–41. *In* J. Strathern, E. Jones, and J. Broach (ed.), The molecular biology of the yeast Saccharomyces. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
16. **Faye, G., K. Leung, K. Tatchell, B. Hall, and M. Smith.** 1981. Deletion mapping of sequences essential for in vivo transcription of the iso-1-cytochrome c gene. Proc. Natl. Acad. Sci. U.S.A. **78:**2258–2262.
17. **Federoff, H., T. Eccleshall, and J. Marmur.** 1983. The regulation of maltase synthesis in *Saccharomyces carlsbergensis*. J. Bacteriol. **154:**1301–1308.
18. **Fromm, M., and P. Berg.** 1983. Transcription *in vivo* from SV40 early promoter: deletion mutants without repression by large T antigen. J. Mol. Appl. Genet. **2:**127–135.
19. **Fukasawa, T., K. Obenai, T. Segawa, and Y. Nogi.** 1980. The enzymes of the galactose cluster in *Saccharomyces cerevisiae*. II. Purification and characterization of uridine diphosphoglucose 4-epimerase. J. Biol. Chem. **255:**2705–2707.
20. **Guarente, L., G. Lauer, T. Roberts, and M. Ptashne.** 1980. Improved methods for maximizing expression of a cloned gene: a bacterium that synthesizes rabbit β-globin. Cell **20:**543–553.
21. **Guarente, L., and T. Mason.** 1983. Heme regulates transcription of the CYC1 gene of *S. cerevisiae* via an upstream sequence. Cell **32:**1279–2386.
22. **Guarente, L., and M. Ptashne.** 1981. Fusion of *Escherichia coli lacZ* to the cytochrome c gene of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U.S.A. **78:**2199–2203.
23. **Guarente, L., R. Yocum, and P. Gifford.** 1982. A *GAL10-CYC1* hybrid yeast promoter identifies the *GAL4* regulator as an upstream site. Proc. Natl. Acad. Sci. U.S.A. **79:**7410–7414.
24. **Hashimoto, H., Y. Kikuchi, Y. Nogi, and T. Fukasawa.** 1983. Regulation of expression of the galactose gene cluster in *Saccharomyces cerevisiae*: isolation and characterization of the regulatory gene *GAL4*. Mol. Gen. Genet. **191:**31–38.
25. **Hopper, J., J. Broach, and L. Rowe.** 1978. Regulation of the galactose pathway in *Saccharomyces cerevisiae*: induction of uridyl transferase mRNA and dependency on *GAL4* gene function. Proc. Natl. Acad. Sci. U.S.A. **75:**2878–2882.
26. **Johnson, A., A. Poteete, G. Lauer, R. Sauer, G. Ackers, and M. Ptashne.** 1981. λ repressor and cro—components of an efficient molecular switch. Nature (London) **294:**217–223.
27. **Johnston, M., and R. Davis.** 1984. Sequences that regulate the divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4:**1440–1448.
28. **Johnston, S., and J. Hopper.** 1982. Isolation of the yeast regulatory gene *GAL4* and analysis of its dosage effects on the galactose/melibiose regulon. Proc. Natl. Acad. Sci. U.S.A. **79:**6971–6975.
29. **Kew, O., and H. Douglas.** 1976. Genetic co-regulation of galactose and melibiose utilization in *Saccharomyces cerevisiae*. J. Bacteriol. **125:**33–41.
30. **Klar, A., and H. Halvorson.** 1976. Effect of *GAL4* gene dosage on the level of galactose catabolic enzymes in *Saccharomyces cerevisiae*. Mol. Gen. Genet. **135:**203–211.
31. **Laughon, A., and R. Gesteland.** 1982. Isolation and preliminary characterization of the *GAL4* gene, a positive regulator of transcription in yeast. Proc. Natl. Acad. Sci. U.S.A. **79:**6827–6831.
31a. **Laughon, A., and R. Gesteland.** 1984. Primary structure of the *Saccharomyces cerevisiae GAL4* gene. Mol. Cell. Biol. **4:**260–267.
32. **Leloir, L.** 1951. The enzymatic transformation of uridine diphosphate glucose into a galactose derivative. Arch. Biochem. **33:**186–190.
33. **Mandelstam, J.** 1962. The repression of constitutive β-galactosidase by glucose and other carbon sources. Biochem. J. **84:**489–493.
34. **Maniatis, T., E. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
35. **Matsumoto, K., Y. Adachi, A. Toh-e, and Y. Oshima.** 1980. Function of positive regulatory gene *GAL4* in the synthesis of galactose pathway enzymes in *Saccharomyces cerevisiae*: evidence that the *GAL81* region codes for part of the *GAL4* protein. J. Bacteriol. **141:**508–527.
36. **Matsumoto, K., A. Toh-e, and Y. Oshima.** 1978. Genetic control of galactokinase synthesis in *Saccharomyces cerevisiae*: evidence for constitutive expression of the positive regulatory gene *GAL4*. J. Bacteriol. **134:**446–457.
37. **Matsumoto, K., A. Toh-e, and Y. Oshima.** 1981. Isolation and characterization of dominant mutations resistant to carbon catabolite repression of galactokinase synthesis in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **1:**83–93.
38. **Matsumoto, K., I. Uno, A. Toh-e, T. Ishikawa, and Y. Oshima.** 1982. Cyclic AMP may not be involved in catabolite repression in *Saccharomyces cerevisiae*: evidence from mutants capable of utilizing it as an adenine source. J. Bacteriol. **150:**277–285.
39. **Matsumoto, K., T. Yoshimatsu, and Y. Oshima.** 1983. Recessive mutations conferring resistance to carbon catabolite repression of galactokinase synthesis in *Saccharomyces cerevisiae*. J. Bacteriol. **153:**1405–1414.
40. **Maxam, A., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74:**560–564.
41. **McKnight, S., and R. Kingsbury.** 1982. Transcriptional control signals of a eukaryotic protein-coding gene. Science **217:**316–324.
42. **Miller, J.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
43. **Nogi, Y., and T. Fukasawa.** 1980. A novel mutation that affects utilization of galactose in *Saccharomyces cerevisiae*. Curr. Genet. **2:**115–120.
44. **Nogi, Y., and T. Fukasawa.** 1983. Nucleotide sequence of the transcriptional initiation region of the yeast *GAL7* gene. Nucleic Acids Res. **11:**8555–8568.
45. **Nogi, Y., K. Matsumoto, A. Toh-e, and Y. Oshima.** 1977. Interaction of super-repressible and dominant constitutive mutations for the synthesis of galactose pathway enzymes in *Saccharomyces cerevisiae*. Mol. Gen. Genet. **152:**137–144.
46. **Orr-Weaver, T., J. Szostak, and R. Rothstein.** 1981. Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. U.S.A. **78:**6354–6358.
47. **Oshima, Y.** 1982. Regulatory circuits for gene expression: the metabolism of galactose and phosphate, p. 159–180. *In* J. Strathern, E. Jones, and J. Broach (ed.), The molecular biology of the yeast Saccharomyces. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
48. **Perlman, D., and J. Hopper.** 1979. Constitutive synthesis of the *GAL4* protein, a galactose pathway regulator in *Saccharomyces cerevisiae*. Cell **16:**89–95.
49. **Rose, M., M. Casadaban, and D. Botstein.** 1981. Yeast genes fused to β-galactosidase in *Escherichia coli* can be expressed normally in yeast. Proc. Natl. Acad. Sci. U.S.A. **78:**2460–2464.
50. **Russell, P.** 1983. Evolutionary divergence of the mRNA transcription initiation mechanism in yeast. Nature (London) **301:**167–169.
51. **Schell, M., and D. Wilson.** 1979. Cloning and expression of the *Saccharomyces cerevisiae* galactokinase gene in an *Escherichia coli* plasmid. Gene **5:**291–303.

52. **Schlesinger, D., M. Schell, and D. Wilson.** 1977. The NH_2-terminal sequences of galactokinase from *Escherichia coli* and *Saccharomyces cerevisiae*. FEBS Lett. **83:**45–47.
53. **Segawa, T., and T. Fukasawa.** 1979. The enzymes of the galactose cluster in *Saccharomyces cerevisiae*. Purification and characterization of galactose-1-phosphate uridyl transferase. J. Biol. Chem. **254:**10,717–10,709.
54. **Segawa, T., and T. Fukasawa.** 1980. Transcriptional units of *GAL* genes in *Saccharomyces cerevisiae* determined by ultraviolet light mapping. Curr. Genet. **2:**223–228.
55. **Sentenac, A., and B. Hall.** 1982. Yeast nuclear RNA polymerases and their role in transcription, p. 561–606. *In* J. Strathern, E. Jones, and J. Broach (ed.), The molecular biology of the yeast Saccharomyces. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
56. **Sherman, F., G. Fink, and J. Hicks.** 1981. Methods in yeast genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
57. **Sherman, F., J. Stewart, and A. Schweingruber.** 1980. Mutants of yeast initiating translation of iso-1-cytochrome C within a region spanning 37 nucleotides. Cell **20:**215–222.
58. **Stinchcomb, D., C. Mann, and R. Davis.** 1982. Centromeric DNA from *Saccharomyces cerevisiae*. J. Mol. Biol. **158:**157–179.
59. **St. John, T., and R. Davis.** 1979. Isolation of galactose inducible sequences from *Saccharomyces cerevisiae* by differential plaque filter hybridization. Cell **16:**443–452.
60. **St. John, T., and R. Davis.** 1981. The organization and transcription of the galactose gene cluster of *Saccharomyces*. J. Mol. Biol. **152:**285–315.
61. **St. John, T., S. Scherer, M. McDonnell, and R. Davis.** 1981. Deletion analysis of the *GAL* gene cluster: transcription from three promoters. J. Mol. Biol. **152:**317–337.
62. **Struhl, K.** 1982. The HIS3 promoter contains at least two distinct elements. Proc. Natl. Acad. Sci. U.S.A. **79:**7385–7389.