MOLECULAR AND CELLULAR BIOLOGY, Feb. 1984, p. 268–275
0270-7306/84/020268-08$02.00/0

Vol. 4, No. 2

Identification of Two Proteins Encoded by the *Saccharomyces cerevisiae GAL4* Gene

ALLEN LAUGHON,† ROBERT DRISCOLL, NORMA WILLS, AND RAYMOND F. GESTELAND*

Department of Biology and Howard Hughes Medical Institute, University of Utah, Salt Lake City, Utah 84112

Received 28 July 1983/Accepted 9 September 1983

We placed the *Saccharomyces cerevisiae GAL4* gene under control of the galactose regulatory system by fusing it to the *S. cerevisiae GAL1* promoter. After induction with galactose, *GAL4* is now transcribed at about 1,000-fold higher levels than in wild-type *S. cerevisiae*. This regulated high-level expression has enabled us to tentatively identify two *GAL4*-encoded proteins.

In *Saccharomyces cerevisiae*, the synthesis of a set of enzymes involved in galactose utilization is under the control of a well-defined regulatory system. These galactose-inducible proteins, galactokinase, galactose permease, α-D-galactose-1-phosphate uridyltransferase, uridine diphosphogalactose-4-epimerase, and α-galactosidase are encoded by the *GAL1*, *GAL2*, *GAL7*, *GAL10*, and *MEL1* genes, respectively (8, 18). The *GAL1*, *GAL7*, and *GAL10* genes are clustered within a 7-kilobase (kb) region of DNA on chromosome II, and it has been shown that galactose regulation of these genes is at the transcriptional level (16, 33).

The *GAL4* and *GAL80* genes encode proteins which are required for the regulation of the galactose-inducible genes (8, 9). *GAL4* encodes a positive regulator of transcription which in the case of the *GAL10* gene acts at a site 130 to 500 base pairs (bp) upstream from where transcription starts (15). *GAL80* encodes a negative regulator which, in the absence of galactose, interacts with the *GAL4* product to prevent the activation of transcription by *GAL4* (25, 26, 30). Although the *GAL1*, *GAL7*, and *GAL10* gene cluster has been cloned and characterized (33–35), almost nothing is known about the interaction of the *GAL4* product with DNA regulatory sites or with the *GAL80* product.

Recently, the *GAL4* gene has been cloned (17, 21), with one aim being the characterization of its encoded product. It was found that the *GAL4* transcript is present constitutively at very low levels (0.1 transcript per cell), such that characterization of the encoded protein presents serious difficulties (21). The DNA sequence of the gene showed that *GAL4* was capable of encoding a 99,350-dalton protein (22). Mapping of *GAL4* transcripts with S1 nuclease revealed that the 5′ ends are heterogeneous, with 70% of the transcripts starting 10 to 20 bp upstream and 30% starting downstream from the AUG where translation of the 99,350-dalton protein would begin. It was reasoned that in these shorter transcripts translation may initiate at a downstream AUG, resulting in the synthesis of a 91,600-dalton protein lacking the first 78 amino acids of the larger protein. What role the two postulated *GAL4* proteins might play in activation of galactose-inducible genes is unclear.

It seemed likely that it would be necessary to "overproduce" the *GAL4* proteins before they could be characterized directly. We report here the fusion of *GAL4* to the yeast *GAL1* promoter, which results in regulated, high-level transcription of *GAL4*. This approach has allowed us to identify two *GAL4* proteins as products of in vitro translation programmed by *GAL4* mRNA.

* Corresponding author.

† Present address: Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309.

MATERIALS AND METHODS

Strains and media. *S. cerevisiae* strains S32A (α *leu2-3 leu2-112 gal4-4*) and S61C (α *leu2-3 leu2-112*) were obtained by sporulation and tetrad dissection of a diploid from the mating of strains DC5 (**a** *leu2-3 leu2-112 his3 can1*) and GAL4-4 (α *gal4-4 leu1-2 his5-2*). Strains DC5 and GAL4-4 were gifts from J. Broach. S61C strains, transformed with *GAL1/GAL4* fusion plasmids, were named according to the plasmid: pGF1/4, strain GF1/4; p90-1, strain 90-1; etc. Yeast strains were grown at 30°C in either YEP (2% peptone–1% yeast extract) or YNB (0.67% yeast nitrogen base plus amino acids), and 2% glucose, 2% galactose, or 3% glycerol (vol/vol) was provided as a carbon source, with 2% agar for plates. In addition, galactose indicator plates contained 0.003% bromothymol blue. Galactose induction of S61C and *GAL1/GAL4* fusion strains was done as follows: A flask of YNB minus leucine with 2% glucose was inoculated from a colony on a plate and grown to stationary phase. The culture was then diluted 1:20 into a flask of YEP with 3% glycerol and grown for 10 h before a 1/10 volume of 20% galactose was added.

Construction of plasmids. Purification of plasmid DNA and of restriction fragments from gels and transformation of *Escherichia coli* and *S. cerevisiae* with plasmid DNA was performed as described previously (22). The standard procedures which were used in the enzymatic manipulation of cloned DNA are described by Maniatis et al. (23). Plasmid pGF1/4 (Fig. 1) was constructed in the following steps: (i) pG525 was cleaved with *Sph*I and *Acc*I; the *Acc*I sites were filled in by using avian myeloblastosis virus reverse transcriptase; and the 85-bp *Acc*I/*Sph*I fragment spanning the *GAL4* transcriptional start sites was ligated into *Bam*HI/*Sph*I-cleaved YEp13 after the *Bam*HI sticky end was filled in. (ii) The resulting plasmid, pBS1 (containing a regenerated *Bam*HI site at the *Bam*HI/*Acc*I junction), was cleaved with *Hin*dIII, and the sticky ends were filled in, followed by cleavage with *Bam*HI; the 755-bp *GAL1/GAL10* divergent promoter fragment from BM126, with a sticky *Bam*HI end and a filled-in *Eco*RI end, was ligated to the pBS1 DNA. (iii) The resulting plasmid, pF14, was cleaved with *Sph*I and ligated to the 3.2-kb *Sph*I fragment of pG525, which extends from the 5′ end of *GAL4* through to the tet^r gene of pBR322. (iv) A resulting clone with the intact *GAL4* gene fused to the *GAL1* promoter was designated pGF1/4.

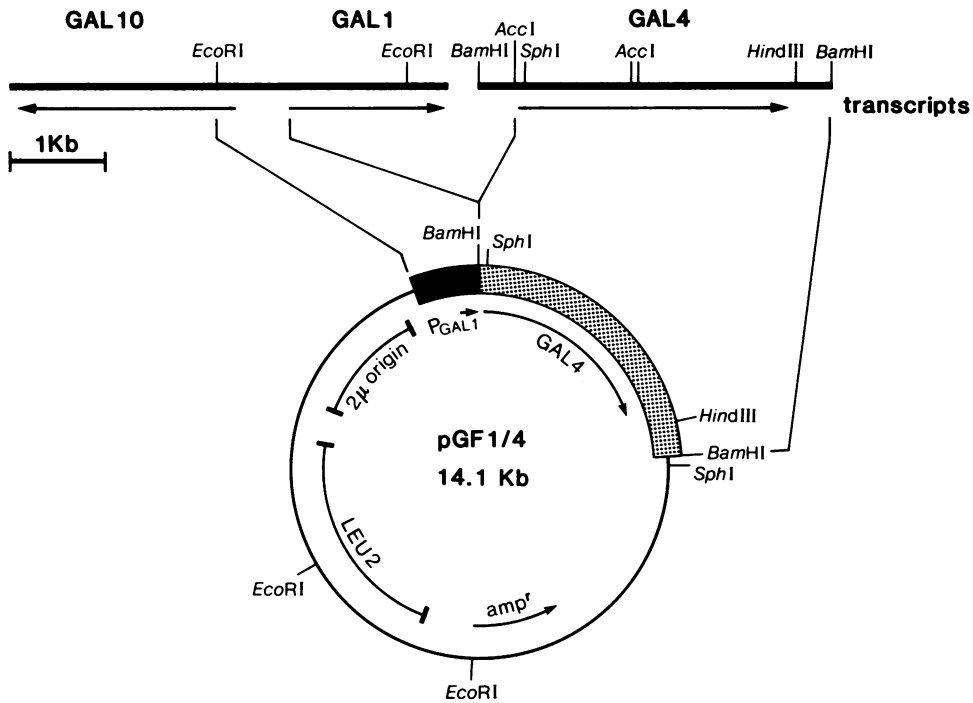

FIG. 1. The structure of the *GAL1/GAL4* fusion plasmid pGF1/4. Plasmid pGF1/4 is YEp13 with a *GAL1/GAL4* fusion substituted for the region between the *Hin*dIII site adjacent to tet^r (the *Hin*dIII site was lost during the construction of pGF1/4) and the *Bam*HI site within tet^r. The *GAL4* segment extends from an *Acc*I site 40 bp upstream of the upstream cluster of *GAL4* transcription starts to a *Bam*HI site in pG525, 500 bp downstream of *GAL4*. The *GAL1* segment extends from a *Bam*HI linker positioned at the *GAL1* transcriptional start to an *Eco*RI site in the *GAL10* gene which was lost during the construction of pGF1/4. A *Bam*HI site was regenerated at the *GAL1/GAL4* junction from the fusion of filled-in *Bam*HI and *Acc*I sites.

Plasmids p90-1, p90-4, and p90-6 were generated in the following steps: (i) pF14 was cleaved with *Bam*HI and digested with *Bal*31 at 20 μg of DNA per ml and 6 U of *Bal*31 per ml at 30°C for 90 s; recessed 3′ ends were filled in with the Klenow fragment of DNA polymerase I, and the DNA was ligated. (ii) The resulting plasmids were screened for deletion size by mapping with restriction endonucleases, and several with deletions of 20 to 60 bp were reconstructed by ligating in the 3.2-kb *Sph*I fragment of pG525 as described for pGF1/4.

RNA purification. RNA was purified from yeast cells according to the method of Elder et al. (10), except that diethyl pyrocarbonate was not used during the procedure. Polyadenylated RNAs were enriched by polyuridylate-Sepharose chromatography as described by St. John and Davis (33). tRNA for use in in vitro translations was purified from crude yeast RNA by binding RNA to a Whatman DE52 DEAE-cellulose column in 20 mM Tris (pH 7.5), followed by repeated washing of the column with the same buffer and elution of the tRNA in 1 M NaCl–20 mM Tris (pH 7.5).

Transfer of RNA to DBM paper and hybridization. RNAs were fractionated by electrophoresis in 1.2% agarose methylmercury hydroxide gels and transferred to diazobenzyloxymethyl paper (DBM paper) as described by Alwine et al. (2). DNA to be used as hybridization probe was labeled with [α-^{32}P]dCTP by nick translation to a specific activity of 5 × 10^8 cpm/μg of DNA according to the method of Maniatis et al. (24).

S1 mapping. S1 nuclease mapping of the 5′ and 3′ ends of the *GAL1/GAL4* fusion transcripts was performed according to the method of Nasmyth et al. (28), as described previously for the natural *GAL4* transcripts (22). A 563-bp *Hin*fI fragment from pG525, which begins in the *GAL4* coding sequence and extends well past the 5′ end of the transcribed region, was 5′-end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. The strands were separated and the mRNA-complementary strand was used in hybridization to RNA before S1 nuclease digestion. Each 100-μl hybridization reaction contained 25 μg of RNA and 5 × 10^5 cpm of DNA. Hybrids protected from S1 nuclease were fractionated on an 8% polyacrylamide sequencing gel, along with a sequencing ladder of the same single-stranded *Hin*FI fragment (28).

Hybrid selection of RNA. Hybrid selection was done following the procedure of Reed et al. (32). pG525 or pBR322 DNA (50 μg of DNA per filter) was coupled to 1-cm^2 DBM paper filters. Hybridization reactions of 100 μl each contained 15 DNA-coupled filters and 300 μg of polyadenylate-enriched RNA purified from strain GF1/4 30 min after a shift to galactose as described above. After hybridization and washing, bound RNA was eluted by incubating the filters in 3 ml of 1 mM EDTA (pH 8) for 10 min in 5°C steps from 45 to 70°C. RNA from each step elution was concentrated by chromatography on a 0.1-ml polyuridylate-Sepharose column as described above. Four micrograms of tRNA, purified from strain GF1/4 30 min after a shift to galactose, was added to each 400-μl elution sample as carrier. RNA was then precipitated by the addition of 20 μl of 5 M NaCl and 1 ml of 95% ethanol.

In vitro translation. In vitro translations with rabbit reticulocyte lysates were performed according to the procedure of Pelham and Jackson (29) as previously described (12). Lysate (150 μl) was supplemented with 2.5 μl of 1 mM hemin and 0.6 μl of creatine phosphokinase (10 mg/ml). To 100 μl of supplemented lysate was added 10 μl of 2 M potassium acetate, 4 μl of 0.5 M creatine phosphate, 2 μl of 0.1 M magnesium acetate, 2 μl of 0.1 M dithiothreitol, 10 μl of 0.5

mM amino acids minus methionine, 1.5 μl of water, 1.3 μl of 0.1 M $CaCl_2$, and 2.6 μl of micrococcal nuclease (1 mg/ml). This mixture was incubated at 20°C for 10 min and placed on ice, followed by the addition of 2.6 μl of 0.1 M ethyleneglycol-bis(β-aminoethyl ether)-*N,N*-tetraacetic acid, 20 μl of [^{35}S]methionine (10 mCi/ml), and 4 μl of tRNA (4 mg/ml; purified from yeast GF1/4 30 min after a shift to galactose). Ten microliters of the reaction mixture was added to the RNA samples, and the total volume was adjusted to 12.5 μl with water. Protein synthesis was allowed to proceed for 1 h at 30°C, followed by the addition of 2 μl of 1 mg of RNase A per ml–200 mM EDTA and incubation for another 10 min at 30°C. Fifty microliters of 0.1 M dithiothreitol–2% sodium dodecyl sulfate (SDS)–80 mM Tris (pH 6.9)–10% glycerol–0.004% bromphenol blue was added to each reaction, and the samples were boiled for 3 min and loaded on 10% polyacrylamide-SDS gels (20).

RESULTS AND DISCUSSION

Fusion of *GAL4* to the *GAL1* promoter. The *GAL4* gene is normally transcribed constitutively at about 0.1 copy per cell, which corresponds to about 10^{-5} of the total mRNA population (21). Transcription of the *GAL1* gene is regulated by both galactose and glucose. Growth on galactose results in high-level transcription of *GAL1* (0.25 to 1% of total mRNA), whereas growth on glucose represses this transcription at least 1,000-fold (34). The *GAL1* promoter therefore would seem to be useful for the purpose of producing the *GAL4* protein(s) since conditions could be adjusted so that there would be either high- or low-level expression of a *GAL1/GAL4* fusion product. High-level expression of *GAL4* might feed back on the *GAL1* promoter since *GAL4* is a positive regulator of *GAL1*. This could result in extremely high levels of transcription from the *GAL1* promoter since it has been shown that expression of galactose-inducible genes is elevated by increased *GAL4* gene dosage (17). The ability to modulate the transcription of such a gene fusion should aid in the identification of the *GAL4* protein(s) by virtue of its presence and absence under inducing and noninducing conditions, respectively. It may also be necessary to modulate *GAL4* transcription if high levels of *GAL4* protein are toxic to yeast cells.

Guarente et al. (15) have shown that the *GAL10* regulatory region (presumed target site for *GAL4*) lies within a 365-bp region beginning 133 bp upstream from the *GAL10* mRNA start site and that this region confers *GAL4/GAL80*-mediated regulation of transcription of *CYC1* when substituted for upstream sequences. In addition, it was found that galactose regulation behaves normally when such fusions are carried on a multicopy plasmid. M. Johnston (personal communication) has constructed a set of *Bal*31 nuclease-generated deletions which extend various distances toward the *GAL1* promoter from a position in the *GAL1* coding sequence. *Bam*HI linkers were positioned at the ends of these deletions, and fusions were made to the yeast *HIS3* gene. Fusions were characterized in which the *HIS3* gene had become galactose inducible (M. Johnston, personal communication).

Using one of the *GAL1* deletions, generously provided by M. Johnston, we have fused *GAL4* to the *GAL1* promoter (Fig. 1). The *GAL1* deletion, in the plasmid BM126, has a *Bam*HI linker located at its 3′ endpoint, which is only a few nucleotides upstream from the position of the *GAL1* mRNA 5′ ends (Fig. 2). The *GAL1* fragment in BM126 extends upstream from this *Bam*HI to an *Eco*RI site 755 bp away in the *GAL10* gene. This *Bam*HI/*Eco*RI fragment contains the entire *GAL1/GAL10* divergent promoter and regulatory region (15; M. Johnston, personal communication). The *GAL1* promoter was fused to *GAL4* at an *Acc*I site which is 35 to 40 bp upstream from the major cluster of *GAL4* mRNA 5′ ends (22). This fusion was constructed in the high-copy-number yeast vector YEp13 (5) and was designated pGF1/4. If transcription were to start at the same position in pGF1/4 (relative to the *GAL1* promoter sequences) as it does in the intact *GAL1* gene, transcription of *GAL4* would begin 35 to 40 bp upstream from where it starts in the intact *GAL4* gene. To vary the distance between the *GAL1* promoter and the *GAL4* coding sequences, we constructed a set of *Bal*31 nuclease-generated deletions which extend various distances in both directions from the *Bam*HI site at the *GAL1/GAL4*

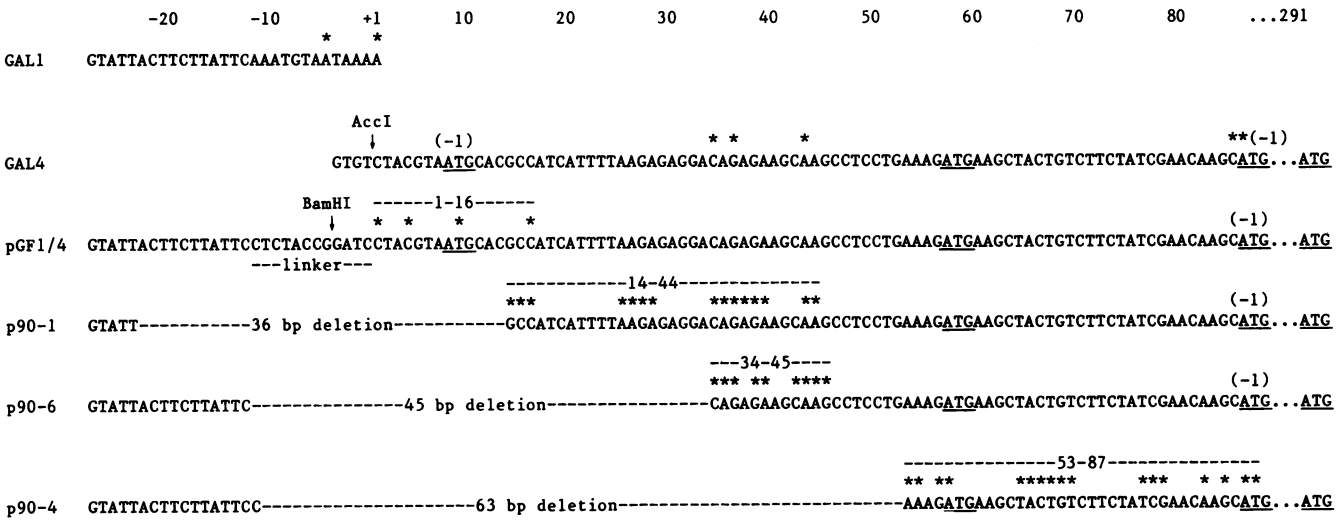

FIG. 2. Junction sequences and transcriptional start sites of *GAL1/GAL4* fusions. Transcriptional starts are indicated by asterisks above the respective sequences. *Bam*HI linker sequences derived from BM126 are indicated below the pGF1/4 sequence. The positions of *GAL1* transcriptional starts and of both natural and BM126-derived *GAL1* sequences are from personal communications from M. Johnston. Putative translation initiation ATGs are at positions 57 and 291 and shifts from the open reading frame are indicated above out-of-frame ATGs.

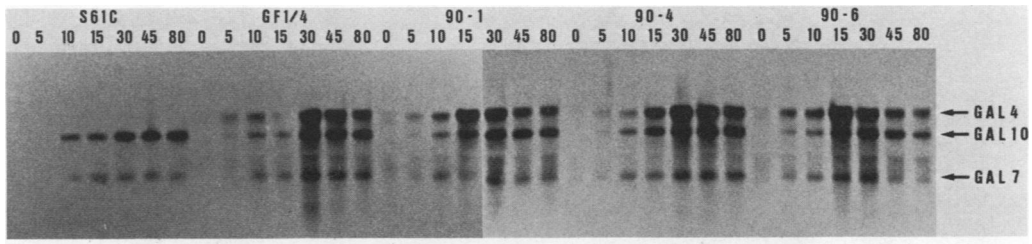

FIG. 3. Kinetics of induction of *GAL1/GAL4* fusions. Five-microgram samples of total RNA, purified from the indicated strains at the indicated number of minutes after a shift to galactose, were electrophoresed in 1.2% agarose methylmercury hydroxide gels, transferred to DBM paper, and hybridized to DNA labeled with [α-^{32}P]dCTP by nick translation. The filters were probed with a 750-bp *Hin*fI fragment internal to the *GAL4* gene and a 5.4-kb *Eco*RI fragment containing the *GAL7* and *GAL10* genes (34). Autoradiographic exposure was for 4 h with intensifying screens. The positions of the 28.8-kb *GAL4*, 2.25-kb *GAL10*, and 1.25-kb *GAL7* mRNAs are indicated.

junction. The deletion plasmids chosen for study, p90-1, p90-4, and p90-6, had lost 36, 63, and 45 bp of DNA respectively (Fig. 2).

Galactose-regulated expression of *GAL1/GAL4* fusions. Plasmids pGF1/4, p90-1, p90-4, and p90-6 were transformed into the *S. cerevisiae* strains S61C (α *leu2-3 leu 2-112*) and S32A (α *leu2-3 leu2-112 gal4-4*), selecting for the *LEU2* marker carried on these plasmids. All four plasmids complemented *gal4*, as determined by growth of the S32A transformants on galactose indicator plates. However, colonies grew slowly on galactose, whereas growth on glucose was normal. In liquid culture, strains S61C and S32A, transformed with the fusion plasmids, grew with a doubling time of about 90 min on glucose but stopped growing for 12 to 24 h when transferred to medium containing galactose. Growth in galactose medium eventually resumed at a slow rate (a 5-h doubling time), but this has not been studied further.

To characterize transcription of the *GAL1/GAL4* fusions, we measured the level of RNA homologous to *GAL4* by fractionation of total cellular RNA on methylmercury hydroxide agarose gels followed by transfer to DBM paper and hybridization with radiolabeled *GAL4* DNA. Initially, we examined RNA from strain S61C, containing pGF1/4 and referred to hereafter as GF1/4, that had been shifted from glucose to galactose. Normal galactose induction did not occur after shifting from glucose to galactose since *GAL7* and *GAL10* mRNAs failed to appear after 5.5 h, although these transcripts were present at slightly lower than wild-type levels 24 h after the shift. Surprisingly, we found no induction of *GAL4* homologous RNA at any time up to 24 h after the shift to galactose (data not shown). It is known, however, that when wild-type cells are switched from glucose to galactose, there is a 2 to 5 h lag before induction of *GAL1*, *GAL7*, and *GAL10* (1), so perhaps this problem is just compounded in the strain carrying plasmid pGF1/4. Galactose induction occurs within 15 to 30 min when wild-type cells are pregrown in media containing a nonfermentable carbon source such as glycerol (4, 34). And when the *GAL1/GAL4* fusion strains were induced by galactose after growth on glycerol, rapid appearence of *GAL4* transcripts followed (Fig. 3). All four of the S61C-derived strains examined (referred to as GF1/4, 90-1, 90-4, and 90-6, according to the plasmids they carry) exhibit this behavior. Levels of *GAL4* transcript peaked at 15 min after the galactose shift for strains 90-1 and 90-6 and after 30 min for strains GF1/4 and 90-4. In each case, the level of *GAL4* transcript declined immediately after it reached its maximum.

The levels of *GAL7* and *GAL10* transcripts in these fusion strains, as determined by hybridization with radiolabeled *GAL7* and *GAL10* DNA (Fig. 3), roughly parallel the rise and fall of *GAL4* RNA after the shift to galactose. Figure 3 also shows that in strain S61C *GAL7* and *GAL10* mRNAs are induced at the same rate as in the fusion strains but remain at fully induced levels thereafter. The *GAL7* and *GAL10* RNAs reach peak levels that are about twofold higher in the fusion strains than in the wild type, presumably due to the elevated level of *GAL4* expression (17, 21). The peak levels of *GAL4* mRNA in the fusion strains are about twofold higher than the *GAL7* and *GAL10* transcripts in these strains, as might be expected since the fusions are contained on high-copy-number plasmids.

Since cells harboring *GAL1/GAL4* fusions grow slowly on galactose, it is possible that the high level of *GAL4* expression is toxic to yeast cells, perhaps by interfering with transcription in general. To test this possibility, we measured the levels of *URA3* mRNA in the fusion strains during induction by hybridizing the RNA-bound DMB filter shown in Fig. 3 to radiolabeled *URA3* DAN (data not shown). Levels of *URA3* mRNA were equivalent and constant during the time course in both strain S61C and the *GAL1/GAL4* fusion strains. Therefore, the detrimental effect on high levels of *GAL4* expression on transcription does not appear to be universal.

One possibility is that higher than normal rates of galactose catabolism cause the induction to be transient, perhaps by some type of feedback mechanism. The *GAL7* and *GAL10* mRNAs reach peak levels only twofold higher in the *GAL1/GAL4* fusion strains than in induced wild-type cells. However, Johnston and Hopper (17) found that twofold increases in galactokinase and transferase in a strain containing *GAL4* on a multicopy plasmid are not associated with slow growth or with transient galactose induction. Therefore, it seems unlikely that an elevated level of galactose catabolism could be the cause of the transient induction kinetics characteristic of these strains. An interesting possibility is that too much *GAL4* product may specifically repress, instead of activate, the transcription of galactose-inducible genes.

As measured by densitometer tracings of different exposures of the autoradiograph in Fig. 3, the peak levels of *GAL4* transcript which appear in *GAL1/GAL4* fusion strains after induction are about three- to fourfold higher than the levels of *GAL7* and *GAL10* mRNA in fully induced S61C cells. Since these mRNAs have been estimated to constitute 0.25 to 1% of the total mRNA population (34), the *GAL4* mRNA is 0.75 to 4% of the mRNA in the fully induced fusion strains. This is about 1,000-fold more *GAL4* mRNA than is present in wild-type cells (21).

Although the *GAL1/GAL4* fusions give rise to 2.8-kb *GAL4* transcripts, the same size as the naturally occurring

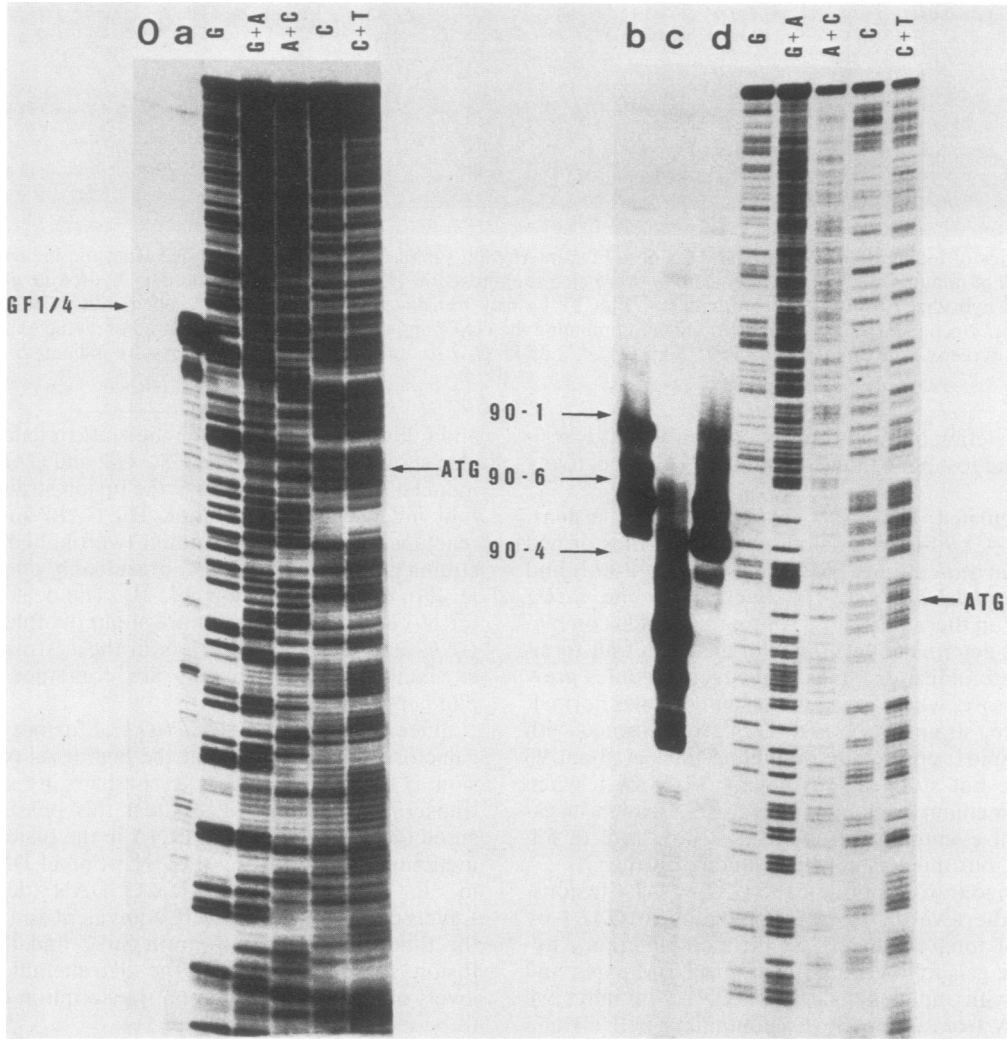

FIG. 4. S1 mapping of the 5′ ends of *GAL1/GAL4* fusion transcripts. The 5′ ends were mapped by protection of a 5′-end-labeled, 563-bp, single-stranded *Hin*fI fragment from S1 nuclease digestion by hybridization to RNA. The protected hybrids were electrophoresed in 8% polyacrylamide sequencing gels, along with sequencing ladders of the same 5′-end-labeled fragment. RNA was from the following strains: lane 0, no RNA; lane a, GF1/4 at 30 min postshift; lane b, 90-1 at 15 min postshift; lane c, 90-4 at 30 min postshift; lane d, 90-6 at 15 min postshift. The positions of the *GAL1/GAL4* junctions for each of the fusions are indicated to the left of the gels, and the position of the AUG which begins the 881-codon *GAL4* open reading frame is to the right. The autoradiographs were exposed for 2 days.

GAL4 mRNA, it was important to determine the position of the 5′ ends of these transcripts, since *GAL4* transcripts normally have heterogeneous 5′ ends and since it has been predicted that this heterogeneity results in the synthesis of two different *GAL4* proteins (22). The 5′ ends of RNAs from the fusion strains were determined by size fractionation on sequencing gels of RNA/DNA hybrids protected from digestion by S1 nuclease (Fig. 4) (3). A 5′-end-labeled *Hin*fI fragment from the intact *GAL4* gene was used to form the RNA/DNA hybrids so that transcripts starting upstream from the *GAL1/GAL4* fusion junctions would appear as bands at the positions of those junctions on the sequencing gels. The positions of the most abundant 5′ ends for each fusion are shown as asterisks above the fusion sequence in Fig. 2. The 5′ ends are heterogeneous in each case and are positioned at an approximately constant distance from a point in the *GAL1* sequence at the left end of Fig. 2. Thus, as is the case with some other promoters (13), RNA polymerase starts transcription at *GAL1* after measuring off a specific distance from an upstream site. In the *GAL1/GAL4* fusions, the positions of the mRNA 5′ ends roughly correspond to the position of the natural *GAL1* mRNA 5′ ends (M. Johnston, personal communication).

It is interesting to note that the 5′ ends of the 90-1 and 90-6 transcripts are clustered around the same positions as the upstream cluster of natural *GAL4* mRNA 5′ ends. This suggests that the position of the 5′ ends is roughly determined by the distance from the promoter but that the exact position is a function of the nucleotide sequence in the immediate vicinity of the 5′ ends. It has been noted previously that either the sequence CAAG or a closely related sequence, possibly important in specifying transcriptional initation in *S. cerevisiae* (6, 7), appears near the 5′ ends of the natural *GAL4* transcripts (22).

GAL4 normally gives rise to transcripts with 5′ ends which fall into two clusters (22), as denoted in Fig. 2. We have proposed that the longer transcripts are translated beginning at the AUG closest to their 5′ ends to produce a 99,350-

dalton protein and that the shorter transcripts are translated beginning at the second AUG from their 5′ ends to yield a 91,600-dalton protein. The major transcripts of strains GF1/4, 90-1, and 90-6 have 5′ ends upstream from initiator AUG of the 99,350-dalton protein, whereas the large majority of the 90-4 transcripts have 5′ ends downstream from this AUG. Given this information, GF1/4, 90-1, and 90-6 strains should produce the larger protein, whereas the 90-4 strain should produce very little of this product. Strain 90-4 might be expected to produce the smaller protein if these mRNAs are efficient templates for translational initiation at the downstream AUG.

Using S1 nuclease mapping, we found that the 3′ ends of the GF1/4 2.8-kb transcripts were identical to those of the natural *GAL4* transcripts (data not shown). The *GAL4* terminator apparently functions efficiently even when transcription is driven from a strong promoter, such as *GAL1*.

At first glance it seems suprising that p90-4 is able to complement a *gal4* mutation, since the bulk of the 90-4 transcripts start downstream from the AUG required for synthesis of the 99,350-dalton protein. However, the low levels of longer mRNA would still be greater than wild-type levels of *GAL4* transcripts and would be expected to be sufficient for complementation.

Identification of *GAL4* proteins. Given the high levels of *GAL4* transcripts in the induced *GAL1/GAL4* fusion strains, it seems likely that *GAL4* protein(s) would be one of the major translation products for some period of time after the shift to galactose. We have not yet been able to identify *GAL4* protein(s) in the *GAL1/GAL4* fusion strains by in vivo labeling with [^{35}S]methionine and fractionation on polyacrylamide-SDS gels. However, galactokinase and UDP-galactose epimerase, the *GAL1* and *GAL10* proteins, appear as major bands on polyacrylamide-SDS gels of crude extracts from strain S61C or the *GAL1/GAL4* fusion strains, pulse labeled with [^{35}S]methionine shortly after the shift from glycerol- to galactose-containing medium (data not shown). Perhaps the *GAL4* mRNAs are being translated poorly, the *GAL4* protein(s) is unstable, or our procedures for extracting proteins are inefficient for the *GAL4* protein(s).

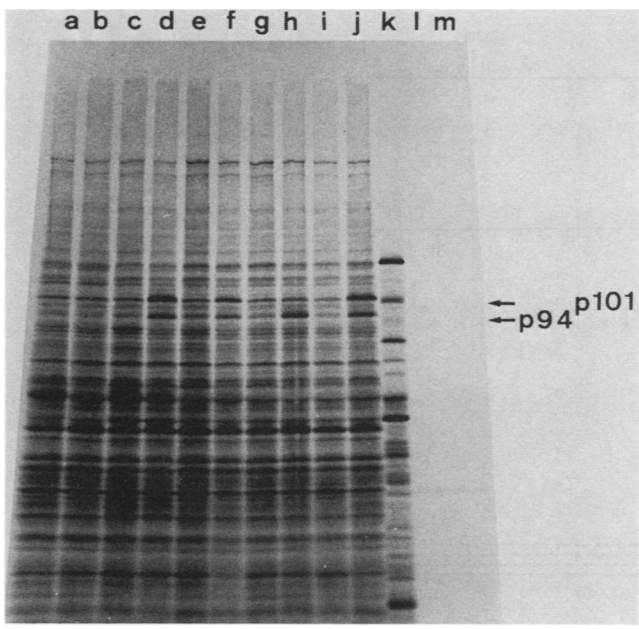

FIG. 5. In vitro translation of RNAs from *GAL1/GAL4* fusion strains. In each case, 5 μg of total RNA was translated in a 12.5-μl reaction mix. After the reactions were stopped and 50 μl of loading buffer was added, 5 μl of each sample was loaded on a 10% polyacrylamide-SDS gel. After electrophoresis, the gel was fixed, dried, and used to expose single-sided X-ray film for 3 days. RNAs were prepared from the following strains: lanes a and b, S61C; lanes c and d, GF1/4; lanes e and f, 90-1; lanes g and h, 90-4; lanes i and j, 90-6. Lane k, Ad2-infected HeLa cells; lane l, no RNA added; lane m, no RNA and no tRNA added to lysate. RNAs were prepared for cells before or at the specified time after a shift to galactose: lanes a, c, e, g, and i, before shift; lanes b, d, and h, 30 min postshift; lanes f and j, 15 min postshift.

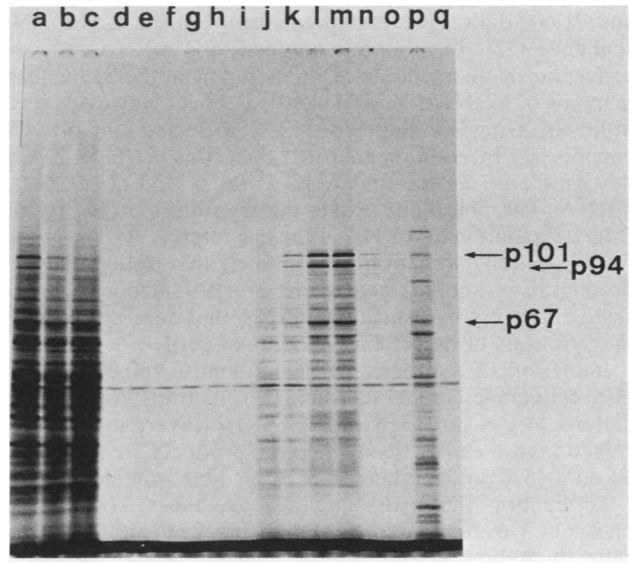

FIG. 6. In vitro translation of hybrid-selected RNAs from galactose-induced *GAL1/GAL4* fusion strain GF1/4. In vitro translation, sample preparation, and electrophoresis were done as described in the legend to Fig. 5, except that 10-μl samples were loaded on the gel. The following amounts of RNAs were used to program protein synthesis in 12.5-μl reactions; lane a, 0.5 μg of polyadenylate-enriched RNA used as starting material for hybrid selection; lane b, 0.5 μg of RNA unbound to pBR322-coupled filters; lane c, 0.5 μg of RNA unbound to pG525-coupled filters; lanes d through i, 1/20 of RNA eluted from pBR322-coupled filters at (lane d) 45°C, (lane e) 50°C, (lane f) 55°C, (lane g) 60°C, (lane h) 65°C, and (lane i) 70°C; lanes j through o, 1/20 of RNA eluted from pG525-coupled filters at (lane j) 45°C, (lane k) 50°C, (lane l) 55°C, (lane m) 60°C, (lane n) 65°C, and (lane o) 70°C; lane p, 2 μg of RNA from Ad2-infected HeLa cells; lane 9, no RNA.

An alternative means of identifying the *GAL4* protein(s) is to use *GAL4* mRNA to program an in vitro protein synthesis system. Preparations of total RNA from strains S61C, GF1/4, 90-1, 90-4, and 90-6 were translated in vitro in a rabbit reticulocyte lysate (29), and the [^{35}S]methionine-labeled products were fractionated on a 10% polyacrylamide-SDS gel (20) (Fig. 5). A 101,000-dalton protein (p101) is a major product with RNA from induced GF1/4, 90-1, and 90-6 cells but is a minor product with RNA from noninduced cells or from induced S61C or 90-4 cells. The apparent low levels of p101 in noninduced cells and induced S61C or 90-4 cells could be due to the presence of another protein migrating with p101. A 94,000-dalton protein (p94) is present with RNA from induced GF1/4, 90-1, 90-4 and 90-6 cells but not with RNA from induced S61C cells. Protein p94 is not made with RNA from noninduced cells.

Proteins p101 and p94 are very close in size to the 99,350-

and 91,600-dalton proteins predicted from the *GAL4* DNA sequence (22). In addition, induced 90-4 mRNA only encodes the smaller protein, as predicted from the S1 nuclease mapping of mRNA 5′ ends described above. However, it is a little surprising that induced GF1/4, 90-1, and 90-6 mRNAs encode p94 in addition to p101 since this AUG is 234 bp downstream from the first AUG at the 5′ end of the *GAL4* mRNA. The possibility exists that synthesis of p94 results from degradation of mRNA, exposing internal AUGs as sites for translational initiation. It is unlikely that p94 is solely the product of proteolytic processing of p101 since it is synthesized efficiently by using 90-4 mRNA that does not program the synthesis of detectable amounts of p101.

Induction of synthesis of galactokinase (60,000 daltons [4]), epimerase (78,000 daltons [11]), or transferase (38,000 daltons [4]) is not seen in Fig. 5. However, induction of galactokinase can be seen when the products are fractionated on a 15% polyacrylamide-SDS gel (data not shown).

To confirm that p101 and p94 are the *GAL4*-encoded products, we enriched for *GAL4* mRNA from induced GF1/4 cells by hybrid selection (31). Plasmid pG525, which is pBR322 with a 3.7-kb yeast DNA insert containing the *GAL4* gene (21), was bound to DBM paper filters (14) and hybridized to polyadenylate-enriched RNA. Unbound RNA was washed off the filters, followed by elution of the hybridized RNA in 1 mM EDTA (pH 8) in 5°C steps from 45 to 70°C. Eluted RNA was translated in vitro as before. As a control, the entire procedure was carried out in parallel with pBR322-coupled DBM paper filters. Figure 6 shows that p101 and p94 are the two major products encoded by the pG525 hybrid-selected mRNA. However, synthesis of a number of less abundant and smaller polypeptides occurs with RNA which elutes at high temperature from the pG525 but not the pBR322 filters. The band at 67,000 daltons could be accounted for by translation of a 1.8-kb polyadenylated transcript known to be homologous to a stretch of pG525 DNA upstream from *GAL4* (21). The other translation products may be due to premature translational termination of *GAL4* mRNA or to weak hybridization of mRNAs that are partially homologous to some region of the 3.7-kb insert of pG525.

A model for *GAL4* transcription and translation. The *GAL4* gene contains an 881-codon open reading frame capable of encoding a 99,350-dalton protein (22). Seventy percent of the *GAL4* mRNAs start at positions upstream from the beginning of this coding region, whereas 30% of the transcripts start downstream of the first AUG in this reading frame and would only be functional if translation initiated at a downstream AUG or if the RNAs were processed. By constructing gene fusions which start transcription at various positions in *GAL4*, we have found that transcripts starting upstream of the first AUG in the 881-codon open reading frame encode both a 101,000-dalton protein and a 94,000-dalton protein, whereas those starting downstream from this AUG encode a 94,000-dalton protein. That the longer transcripts also encode the smaller protein suggests that translation can initiate at either of two AUGs in these messages. Since the 90-6 transcripts start at the same positions as the upstream cluster of natural *GAL4* mRNA starts, it seems likely that natural *GAL4* upstream cluster also gives rise to the two proteins (Fig. 7). Although translation initiates at

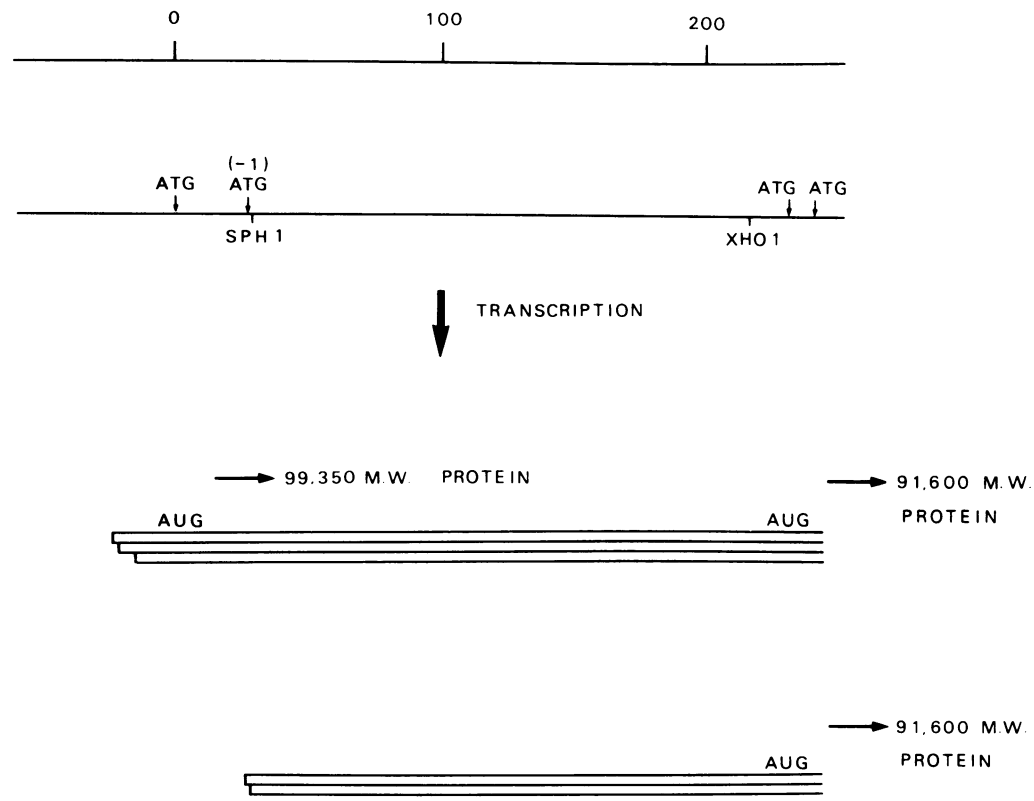

FIG. 7. A model for transcription and translation of *GAL4*. *GAL4* DNA is shown at the top along with the positions of ATGs both in frame and out of frame with the 881-codon open reading frame. Below are long and short *GAL4* transcripts, starting upstream and downstream, respectively, of the AUG which begins the 881-codon open reading frame. We propose that ribosomes initiate at two different AUGs on the long transcripts to yield proteins with predicted molecular weights (M.W.) of 99,350 and 91,600. Short transcripts lacking the upstream AUG give rise to only the smaller protein.

only a single AUG (usually the one closest to the 5′ end [36]) in the vast majority of eucaryotic messages, there are several known examples of initiation at two AUGs on a single messenger (19). However, the possibility exists that the patterns of *GAL4* translational initiation are peculiar to in vitro protein synthesis in the rabbit reticulocyte lysate. Confirmation of this mechanism of *GAL4* expression will require identification of the *GAL4* proteins in vivo.

ACKNOWLEDGMENTS

We are grateful to Mark Johnston for providing us with plasmids and unpublished information on the structure of the *GAL1* promoter. Thanks go to Tom Petes for comments on this manuscript.

A.L. was supported in part by NIH training grant GM07531.

LITERATURE CITED

1. **Adams, B. G.** 1972. Induction of galactokinase in *Saccharomyces cerevisiae*: kinetics of induction and glucose effects. J. Bacteriol. **111**:308–315.
2. **Alwine, J. C., D. J. Kemp, and G. R. Stark.** 1977. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc. Natl. Acad. Sci. U.S.A. **74**:5350–5354.
3. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. Cell **12**:721–732.
4. **Broach, J. R.** 1979. Galactose regulation in *Saccharomyces cerevisiae*: the enzymes encoded by the *GAL7*, 10, 1 cluster are coordinately controlled and separately translated. J. Mol. Biol. **131**:41–53.
5. **Broach, J. R., J. N. Strathern, and J. B. Hicks.** 1979. Transformation in yeast: development of a hybrid cloning vector and isolation of the *CAN1* gene. Gene **8**:121–133.
6. **Burke, R. L., P. Tekamp-Olson, and R. Najarian.** 1983. The isolation, characterization and sequence of the pyruvate kinase gene of *Saccharomyces cerevisiae*. J. Biol. Chem. **258**:2193–2201.
7. **Dobson, M. J., M. F. Tuite, N. A. Roberts, A. J. Kingsman, S. M. Kingsman, R. E. Perkins, S. C. Conroy, B. Dunbar, and L. A. Fothergill.** 1982. Conservation of high efficiency promoter sequences in *Saccharomyces cerevisiae*. Nucleic Acids Res. **10**:2625–2637.
8. **Douglas, H. D., and D. C. Hawthorne.** 1964. Enzymatic expression and genetic linkage of genes controlling galactose utilization in *Saccharomyces*. Genetics **49**:837–844.
9. **Douglas, H. D., and D. C. Hawthorne.** 1966. Regulation of genes controlling synthesis of the galactose pathway enzymes in yeast. Genetics **54**:911–916.
10. **Elder, R. T., E. Y. Loh, and R. W. Davis.** 1983. RNA from yeast transposable element Ty*1* has both ends in the direct repeats, a structure similar to retrovirus RNA. Proc. Natl. Acad. Sci. U.S.A. **80**:2432–2436.
11. **Fukasawa, T., K. Obanai, T. Segawa, and Y. Nogi.** 1980. The enzymes of the galactose cluster in *Saccharomyces cervisiae*. II. Purification and characterization of uridine diphosphoglucose-4-epimerase. J. Biol. Chem. **255**:2705–2707.
12. **Gesteland, R. F., N. Wills, J. B. Lewis, and T. Grodzicker.** 1977. Identification of amber and ochre mutants of the human virus Ad^{2+}NDl. Proc. Natl. Acad. Sci. U.S.A. **74**:4567–4571.
13. **Gluzman, Y., J. Sambrook, and R. J. Frisque.** 1980. Expression of early genes of origin-defective simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **77**:3898–3902.
14. **Goldberg, M. L., R. Lifton, G. R. Stark, and J. G. Williams.** 1979. Isolation of specific RNAs using DNA covalently linked to diazobenzyloxymethyl cellulose or paper. Methods Enzymol. **68**:206–220.
15. **Guarente, L., R. R. Yocum, and P. Gifford.** 1982. A *GAL10-CYC1* hybrid yeast promoter identifies the *GAL4* regulatory region as an upstream site. Proc. Natl. Acad. Sci. U.S.A. **79**:7410–7414.
16. **Hopper, J. E., J. R. Broach, and L. B. Rowe.** 1978. Regulation of the galactose pathway in *Saccharomyces cervisiae*: induction of uridyl transferase mRNA and dependency on *GAL4* gene function. Proc. Natl. Acad. Sci. U.S.A. **75**:2878–2882.
17. **Johnston, S. A., and J. E. 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.
18. **Kew, O. M., and H. C. Douglas.** 1976. Genetic co-regulation of galactose and melibiose utilization in *Saccharomyces*. J. Bacteriol. **125**:33–41.
19. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**:1–45.
20. **Laemmli, U. K.** 1970. Cleavage of structural proteins during assembly of the head of the bacteriophage T4. Nature (London) **227**:680–685.
21. **Laughon, A., and R. F. 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.
22. **Laughon, A., and R. F. Gesteland.** 1984. Primary structure of the *Saccharomyces cerevisiae GAL4* gene. Mol. Cell. Biol. **4**:260–267.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
24. **Maniatis, T., A. Jeffrey, and D. G. Kleid.** 1975. Nucleotide sequence of the righward operator of phage λ. Proc. Natl. Acad. Sci. U.S.A. **72**:1184–1188.
25. **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.
26. **Matsumoto, K., A. Toh-e, and Y. Oshima.** 1978. Genetic control of galactokinase synthesis in *Saccharomyces cervisiae*: evidence for constitutive expression of the positive regulatory gene *gal4*. J. Bacteriol. **134**:446–457.
27. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
28. **Nasmyth, K. A., K. Tatchell, B. D. Hall, C. Astell, and M. Smith.** 1981. A physical analysis of the mating type loci in *Saccharomyces cerevisiae*. Cold Spring Harbor Symp. Quant. Biol. **45**:961–981.
29. **Pelham, H. R. B., and R. Jackson.** 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. **67**:247–256.
30. **Perlman, D., and J. E. Hopper.** 1979. Constitutive synthesis of the *GAL4* protein, a galactose pathway regulator in Saccharomyces cerevisiae. Cell **16**:89–95.
31. **Prives, C. L., H. Aviv, B. M. Paterson, B. E. Roberts, S. Rozenblatt, M. Revel, and E. Winocour.** 1974. Cell free translation of messenger RNA of simian virus 40: synthesis of the major capsid protein. Proc. Natl. Acad. Sci. U.S.A. **71**:302–306.
32. **Reed, S. I., J. Ferguson, and J. C. Groppe.** 1982. Preliminary characterization of the transcriptional and translational products of the *Saccharomyces cerevisiae* cell division cycle gene *CDC28*. Mol. Cell. Biol. **2**:412–425.
33. **St. John, T. P., and R. W. Davis.** 1979. Isolation of galactose-inducible DNA sequence from *Saccharomyces cerevisiae* by differential plaque filter hybridization. Cell **16**:443–452.
34. **St. John, T. P., and R. W. Davis.** 1981. The organization and transcription of the galactose gene cluster of *Saccharomyces*. J. Mol. Biol. **152**:285–315.
35. **St. John, T. P., M. W. McDonnell, and R. W. Davis.** 1981. Deletion analysis of the *Saccharmyces* GAL gene cluster: transcription from three promoters. J. Mol. Biol. **152**:317–334.
36. **Stiles, J. I., J. W. Szostak, A. T. Young, R. Wu, S. Consaul, and F. Sherman.** 1981. DNA sequence of a mutation in the leader region of the yeast iso-1-cytochrome c mRNA. Cell **25**:277–284.