Characterization of *TUP1*, a Mediator of Glucose Repression in *Saccharomyces cerevisiae*

FREDERICK E. WILLIAMS AND ROBERT J. TRUMBLY*

Department of Biochemistry and Molecular Biology, Medical College of Ohio, Toledo, Ohio 43699-0008

The *TUP1* and *CYC8* (=*SSN6*) genes of *Saccharomyces cerevisiae* play a major role in glucose repression. Mutations in either *TUP1* or *CYC8* eliminate or reduce glucose repression of many repressible genes and induce other phenotypes, including flocculence, failure to sporulate, and sterility of *MATα* cells. The *TUP1* gene was isolated in a screen for genes that regulate mating type (V. L. MacKay, Methods Enzymol. 101:325–343, 1983). We found that a 3.5-kb restriction fragment was sufficient for complete complementation of *tup1-100*. The gene was further localized by insertional mutagenesis and RNA mapping. Sequence analysis of 2.9 kb of DNA including *TUP1* revealed only one long open reading frame which predicts a protein of molecular weight 78,221. The predicted protein is rich in serine, threonine, and glutamine. In the carboxyl region there are six repeats of a pattern of about 43 amino acids. This same pattern of conserved residues is seen in the β subunit of transducin and the yeast *CDC4* gene product. Insertion and deletion mutants are viable, with the same range of phenotypes as for point mutants. Deletions of the 3' end of the coding region produced the same mutant phenotypes as did total deletions, suggesting that the C terminus is critical for TUP1 function. Strains with deletions in both the *CYC8* and *TUP1* genes are viable, with phenotypes similar to those of strains with a single deletion. A deletion mutation of *TUP1* was able to suppress the *snf1* mutation block on expression of the *SUC2* gene encoding invertase.

The phenomenon of glucose repression in yeast cells has attracted considerable interest as an example of the coordinate control of a large number of genes by an environmental signal, the external glucose concentration. Wild-type yeast cells, when grown on glucose, have lower levels of many enzymes needed to utilize alternate carbon sources. As the glucose in the medium is exhausted, the synthesis of these enzymes is derepressed. This phenomenon, called carbon catabolite repression or glucose repression, in many cases has been demonstrated to act at the level of transcription.

The pathway mediating glucose repression has remained obscure despite considerable efforts at genetic analysis. Mutants that are insensitive to catabolite repression have been isolated in many laboratories (5, 36, 44). Our laboratory isolated mutations in two genes, *cyc8* and *tup1*, that completely abolish glucose repression of the *SUC2* gene encoding invertase (50).

Mutations in *tup1* induce a wide array of phenotypes and have been isolated in screens for various mutant traits. They were first isolated for their ability to take up dTMP from the growth medium, hence the designation *tup1* (thymidine uptake) (54). *flk1* mutants are extremely flocculent or flaky and are insensitive to catabolite repression of maltase, invertase, and α-methylglucosidase (36, 44). A selection for resistance to UV-induced mutation of *CAN1* to *can1* produced *umr7* mutants (20, 21). A selection for mutants with increased expression of iso-2-cytochrome *c* produced *cyc8* and *cyc9* mutants (34). *amm1* mutants stabilize plasmids containing a defective autonomously replicating sequence element (49). *flk1, umr7, cyc9,* and *amm1* have all been shown to be allelic to *tup1* (21, 34, 44, 49).

tup1 and *cyc8* mutants share many phenotypes, including calcium-dependent flocculation (34, 36, 50), mating-type defects in *MATα* cells (5, 20, 34, 54), nonsporulation of

homozygous diploids (34, 36), and derepression of many enzymes under glucose repression. The diversity of phenotypes may arise from the involvement of *TUP1* and *CYC8* in regulatory pathways distinct from glucose repression or may reflect the interaction of glucose repression with these pathways.

SNF1 (*CCR1, CAT1*) encodes a protein kinase that is necessary for the derepression of glucose-repressible genes (6, 9). Mutant *snf1* strains cannot ferment sucrose or utilize other nonfermentable carbon sources. Some of these effects are suppressed by *cyc8* mutations. Double *snf1 cyc8* mutants are constitutive for invertase synthesis (6). It has been postulated that *CYC8* is a negative regulator of glucose repression and that *SNF1* may act through *CYC8* (6).

Mutations in *TUP1* and *CYC8* produce the same array of phenotypes, suggesting that the two genes function at similar points in regulation. However, their specific roles and the mechanism of glucose repression are still unknown, although the latter seems to involve multiple steps or pathways (11, 27). Possible roles for *TUP1* and *CYC8* as negative regulators could be repression of transcription or interference of activation of transcription. Recently *CYC8* (*SSN6*) was cloned, characterized, and sequenced (37, 51). However, no functional role has yet been identified for the *CYC8* gene product. To further understand the roles of *TUP1* and *CYC8* in catabolite repression, *TUP1* has been characterized and sequenced. This work is described here.

MATERIALS AND METHODS

Strains and media. The *Saccharomyces cerevisiae* strains used (Table 1) are of the S288c genetic background. Bacterial strains JM101 (55) and XL1-Blue (Stratagene) were used as hosts for M13 bacteriophage and plasmids pBS(+) and pBS(−) (Stratagene). Strain JM83 (55) was used as the host for pUC-derived plasmids, and RR1 was used as the host for the other plasmids. Yeast cells were grown on YEPD (1% yeast extract, 2% Bacto-Peptone [Difco], 2% glucose) or on

TABLE 1. Yeast strains

Strain	Genotype
RTY110	*MATa leu2 ura3-1 tup1-100*
RTY185	*MATα ade2 tup1-100*
RTY187	*MATα thr4 trp1 ura3-1*
RTY234	*MATa his4-519 leu2-3 leu2-112 ura3-52*
RTY235	*MATα his4-519 leu2-3 leu2-112 trp1-289 ura3-52*
RTY239	*MATa/α leu2/+ +/thr4 tup1-100/+ +/trp1 ura3-1/ura3-1*
RTY306	*MATa/α +/ade2 leu2/+ tup1-100/tup1-100 ura 3-1/+*
RTY308	*MATa/α +/ade2 his4-519/+ leu2-3/+ leu2-112/+ +/tup1-100 ura 3-52/+*
RTY333	*MATa cat1-42 leu2 trp1 ura3*
RTY335	*MATα his4-519 leu2-3 leu2-112 thr4 trp1 ura3-52*
RTY336	RTY335 *tup1-11::URA3*
RTY338	RTY335 *tup1-7::URA3*
RTY363	*MATa his4-519 leu2-3 leu2-112 ura3-52 cyc8-Δ1::LEU2*
RTY418	*MATα his4-519 leu2-3 leu2-112 trp1-289 ura3-52 tup1-Δ1::TRP1*
RTY420	*MATα his4-519 leu2-3 leu2-112 trp1-289 ura3-52 tup1-Δ2::TRP1*
RTY421	*MATα his4-519 leu2-3 leu2-112 trp1-289 ura3-52 tup1-40::URA3*
RTY438	*MATα his4-519 leu2-3 leu2-112 trp1-289 ura3-52 tup1-Δ1::TRP1 cyc8-Δ1::LEU2*

SD synthetic medium with the appropriate supplements (40). All yeast transformations were performed by the lithium acetate method (18).

Invertase assays. Yeast cells were grown in either YEPD or SD medium with the appropriate supplements overnight at 30°C with aeration. Cells were diluted into YEPD and grown for 5 to 6 h or until the A_{600} value was close to 1.0. Cells were then washed twice in distilled H_2O and resuspended in 1 ml of distilled H_2O. Invertase assays were performed as described previously (50). Invertase activity is expressed as units per 10^9 cells, where units are micromoles of glucose produced per minute.

Plasmids. The *TUP1* gene, isolated from a yeast genomic DNA bank in the shuttle vector YEp13, was provided by Vivian MacKay (23). The isolated plasmid, pMJ1, had approximately 9.7 kb of insert DNA. A 3.7-kb portion of this original isolate was subcloned into the *Bam*HI and *Sal*I sites of the YIp5 vector (4) to produce pRT73. This 3.7-kb subcloned DNA contained a 276-bp *Bam*HI-*Sal*I fragment from YEp13 and yeast genomic DNA. pMJ1 was deleted by a 3.3-kb *Hin*dIII fragment and a 1.7-kb *Bam*HI fragment to produce pFW1-1 and pFW1-2, respectively. The former was used for all subsequent subclonings. A 3.5-kb *Hin*dIII-*Pst*I fragment containing *TUP1* was subcloned into the pUC-derived plasmid pOM2 (28) to create plasmid pTXL6. The *TUP1* gene was rescued from this subcloning as a *Xho*I-*Sph*I fragment, which was ligated into the *Sal*I and *Sph*I sites of YEp24 to produce pFW28 and into the *Sal*I and *Sph*I sites of pRT81 (51) to yield pTXL63.

To create insertion mutants and define the boundaries of the gene, a 6.2-kb *Eco*RI-*Sal*I fragment containing *TUP1* was ligated into the *Eco*RI and *Sal*I sites of pBS(+) (Stratagene) to create pFW2. A 6.2-kb *Hin*dIII-*Eco*RI fragment from this subclone was ligated into the *Hin*dIII and *Eco*RI sites of pHSS6 (38) to produce pFW9.

For construction of specific deletions and insertions, a *Hin*dIII fragment from pFW25 containing all of *TUP1* was ligated into pUC1813 (19) cut with *Hin*dIII to produce pFW35. The *TRP1* gene, on an 852-bp *Eco*RI-*Bgl*II fragment

from YRp7, replaced a 3.1-kb *Eco*RI-*Bam*HI fragment or a 378-bp *Eco*RI-*Bgl*II fragment in pFW35 to yield pFW36 and pFW37. The *URA3* gene was subcloned on a 1.1-kb *Hin*dIII fragment from YEp24 into pUC1318 (19) to create pRT86. *URA3* was rescued from this subclone on an *Eco*RI fragment and ligated with pFW35 cut with *Eco*RI to produce pFW40.

For construction of *snf1* deletions, the *SNF1* gene was isolated on a plasmid (pES1) by complementation of the *cat1-42* mutation in RTY333. The presence of the *SNF1* gene on pES1 was confirmed by restriction digests and DNA sequencing. A 3.8-kb *Eco*RI fragment from pES1 was subcloned into pUC1318 (19) to make pRT87. The *Pst*I-*Bcl*I segment of pRT87 comprising most of the coding region of *SNF1* was replaced by a 2.1-kb *Nsi*I-*Bam*HI fragment from YIp5 carrying the *URA3* gene to make pRT125. Yeast strains RTY235 (wild type) and RTY418 (*tup1-Δ1::TRP1*) were transformed with pRT125, giving rise to strains RTY469 (*snf1-Δ100::URA3*) and RTY508 (*tup1-Δ1::TRP1 snf1-Δ100::URA3*).

To show the effects of single and multiple copies of the *TUP1* and *CYC8* genes on complementation of deletions, the *TUP1* gene was subcloned on a 3.5-kb *Hin*dIII-*Pst*I fragment from pTXL6 into pRS315 (41) to create pFW45. The *CYC8* gene was subcloned on a 5.3-kb *Xba*I-*Sal*I fragment into pRS316 (41) to produce pFW46.

DNA sequencing. Restriction fragments from pFW1-1 that corresponded to the area of the *TUP1* gene mapped genetically were subcloned into bacteriophages M13mp18 and M13mp19 (55) and the phagemid vector pBS. Sequencing was performed by using single-stranded bacteriophage DNA or double-stranded plasmid DNA as a template (7). All sequencing reactions were done by the dideoxy-chain termination method, using the Klenow fragment of DNA polymerase I (35) or modified T7 bacteriophage DNA polymerase (Sequenase) (48). Overlapping deletion subclones were created by using the directed deletion protocol of Henikoff with exonuclease III and S1 nuclease (16).

Isolation of RNA and Northern (RNA) blotting. Yeast strains were grown at 30°C in YEPD medium and were harvested in log phase. The cells were collected by centrifugation at 3,500 × *g* for 5 min, washed in LETS buffer (0.1 M LiCl, 1 mM EDTA, 0.1 M Tris hydrochloride, pH 7.5), and recentrifuged. The cells were resuspended in 250 μl of LETS buffer and vortexed with glass beads. The homogenate was extracted a minimum of five times with water-saturated phenol-chloroform-isoamyl alcohol (25:24:1) and two times with chloroform-isoamyl alcohol (24:1). The nucleic acids were precipitated overnight at −20°C by addition of 2 volumes of ethanol, washed with 70% ethanol, and dried. The dried pellets were dissolved in sterile water. Poly(A)$^+$ RNA was selected on an oligo(dT) column (2). RNA was run on a 0.8% formaldehyde-agarose gel in the presence of MOPS (*N,N,N*, morpholinopropanesulfonic acid) buffer, pH 7.5 (10). The RNA was then transferred to nitrocellulose by the method of Southern (43). The blot was probed with a restriction fragment known to be in the coding region which had been amplified and labeled by the method of random primers (12). The amount of RNA loaded on the gel was estimated by A_{260} measurements and confirmed by ethidium bromide staining of the gel.

Primer extension mapping of the *TUP1* mRNA. The 5' end of the *TUP1* mRNA was determined by primer extension, using a synthetic oligonucleotide primer, 5'-CGAAACGC TGGCAGTCA-3', which corresponds to the antisense strand of nucleotides 510 to 526. The primer was purified by electrophoresis on a 15% acrylamide sequencing gel, fol-

lowed by crush-and-soak elution. The primer was labeled at the 5' terminus with [γ-^{32}P]ATP and polynucleotide kinase. The labeled primer (10^5 cpm) was hybridized with 5 μg of poly(A)$^+$ RNA for 6 h at 40°C in hybridization buffer consisting of 40 mM piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES; pH 6.4), 1 mM EDTA, and 0.4 M NaCl. Primer extension reactions were carried out as described previously (1). The reaction products were run on a 6% acrylamide–8 M urea denaturing gel.

Mating efficiencies of tup1 deletion strains. Quantitative mating assays were done as described by Lemontt et al. (21). Cells were grown overnight in YEPD medium and diluted to an optical density of 10^6 cells per ml. Tester strain RTY5 (MATa ade2) was diluted to 10^8 cells per ml. A 1-ml sample of each test strain was mixed with 1 ml of the tester strain in 5 mM EDTA to disperse flocculent cells. The mixtures were then filtered onto 0.45-μm-pore-size nitrocellulose filters, which were transferred to a YEPD plate and incubated for 6 h at 30°C. Cells were resuspended in 1 ml of sterile water. Dilutions were plated on minimal plates to select for diploids and on SD-adenine to determine the number of test strain cells.

Nucleotide sequence accession number. The sequence reported has been submitted to GenBank under accession number M31733.

RESULTS

Localization of the TUP1 gene. Plasmid pMJ1, consisting of 9.7 kb of yeast genomic DNA including the TUP1 gene in the yeast vector YEp13, was obtained from Vivian MacKay (23). We confirmed the presence of the TUP1 gene on pMJ1 by complementation of tup1 mutations and genetic mapping. RTY110 (MATa leu2 ura3-1 tup1-100), when transformed with pMJ1, no longer exhibited the tup1 phenotypes of flocculence and constitutive invertase synthesis. The transformed strain was crossed with RTY185 (MATα ade2 tup1-100), and the resulting diploid sporulated at normal levels, demonstrating that the plasmid could restore sporulation competence to a tup1 homozygous diploid. A 3.7-kb SalI-BamHI fragment from pMJ1 was subcloned into the integrating vector YIp5 to form pRT73, which was used to transform the diploid strain RTY239 to uracil prototrophy. Three individual transformants were allowed to sporulate. Segregation of URA3 and thr4 produced a total of 11 parental ditypes, 9 tetratypes, and no nonparental ditypes. The URA3-thr4 map distance computed from these data {map units = 50[6(nonparental ditypes) + tetratypes]/total} is 22.9 centimorgans, compared with the published tup1-thr4 distance of 22 centimorgans (54). Integration at or near the TUP1 locus supported the conclusion that the TUP1 gene rather than an extragenic suppressor had been cloned.

The TUP1 gene was localized on the 9.7-kb insert DNA by testing the ability of subclones to complement the tup1 phenotypes of constitutive invertase synthesis: flocculence, MATα shmooing, and sporulation deficiency. To facilitate the subclonings, a restriction map of pMJ1 was constructed (Fig. 1). The 6.2-kb HindIII-SalI and 3.5-kb HindIII-PstI fragments fully complemented the tup1-100 mutant phenotypes, while a subclone with a deletion of a 1.7-kb BamHI fragment failed to do so.

The shuttle mutagenesis procedure (38) was used to create insertion mutations and map the endpoints of the gene. Insertions by a minitransposon carrying the URA3 gene into the 6.2-kb EcoRI-SalI fragment are shown in Fig. 2. Restriction analysis revealed 24 different insertions in the DNA. Of
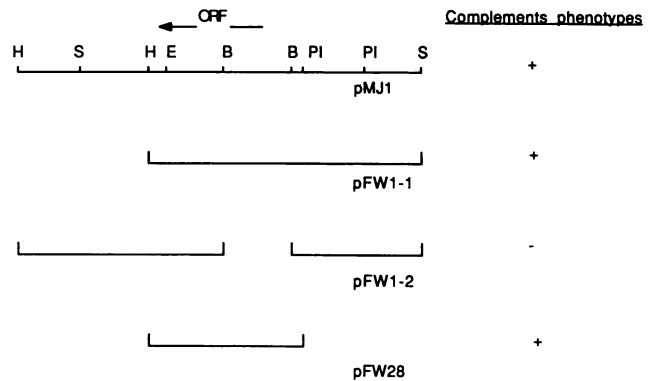

FIG. 1. Localization of the TUP1 gene. Deletions of the pMJ1 insert DNA were tested for complementation of tup1-100 by transformation of RTY110 (MATa leu2 ura3 tup1-100). Complementation was assessed by restoration of glucose repression of invertase and sporulation competence. Restriction sites: H, HindIII; E, EcoRI; S, SalI; PI, PstI; B, BamHI. ORF, Open reading frame revealed by subsequent sequencing; +, complementation of mutant phenotypes; −, no complementation.

these, eight were chosen for transformation of yeast strain RTY335 (MATα his4 leu2 thr4 trp1 ura3) so that insertions could be tested for linkage to thr4. Yeast cells transformed with plasmid DNA digested with NotI were plated on SD-uracil medium to select for uracil prototrophy. Transformants were analyzed for mutant phenotypes of aggregation, shmooing, and constitutive invertase levels. As shown, insertions 11, 6, 7, 14, and 23 induced tup1 mutant phenotypes and were thus in the gene. Insertions 10, 3, and 5 did not induce mutant phenotypes. These results mapped the functional TUP1 gene to a region of approximately 2.6 kb. Insertion mutants were viable and displayed phenotypes similar to those of point mutants. There was a polarity in the strength of the mutant phenotypes, with the most severe phenotypes elicited by insertions near the EcoRI and BglII sites. Surprisingly, the DNA sequence later revealed that these insertions were near the 3' end of the TUP1 open reading frame.

Sequencing of the TUP1 gene. The complete nucleotide sequence of 2.9 kb of DNA including TUP1 was determined on both strands (Fig. 3 and 4). A single long open reading frame of 2,142 nucleotides capable of encoding a protein of molecular weight 78,221 was found in the genetically mapped area. The codon bias index of Bennetzen and Hall (3) identifies genes that are highly expressed and utilize a set of preferred codons. The TUP1 coding region has a codon bias index of 0.22, outside of the class of highly expressed genes, which have values close to 1.0.

The predicted protein has long stretches of glutamine and threonine. Similar stretches of glutamine have been found in many Drosophila proteins regulating embryonic development (32, 52), several yeast regulatory proteins (29, 30, 31, 39), and the rat glucocorticoid receptor (25); however, their functional significance is unknown. Long stretches of threonine have been identified in very few proteins, among them the Drosophila Sgs-3 glue protein (24) and Zipper protein (59). Several regions of the TUP1 protein have high percentages of amino acids which compose the PEST domains proposed to be characteristic of regulatory proteins with short half-lives (33). A search of the TUP1 protein sequence by using the computer program of Rechsteiner (33) identified one domain with a very high PEST score of 12.6, the sequence KETEPENNNTSK (residues 237 to 248).
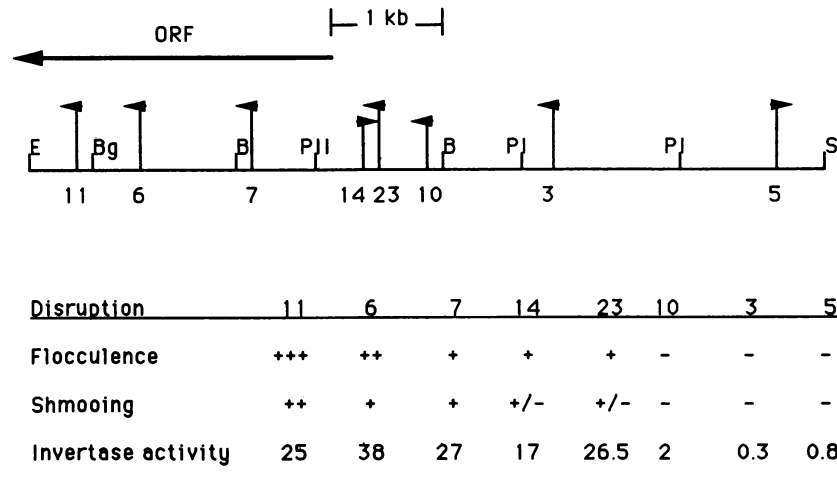
FIG. 2. Transposon mutagenesis. Transposons carrying the ampicillin resistance gene and the yeast *URA3* gene were randomly inserted into the *TUP1* region by the shuttle mutagenesis procedure (37). The 2.4-kb insertions were mapped by restriction analysis. Selected insertions were digested with *Not*I and used to transform RTY335. Points of insertion determined by restriction mapping are denoted by vertical lines. An arrow indicates orientation of the insertion. The severity of the mutant phenotypes is denoted by the number of + signs, from very strong (+ + +) to wild type (−). ORF, Open reading frame deduced from DNA sequence. Restriction sites: E, *Eco*RI; B, *Bam*HI; Bg, *Bgl*II; PI, *Pst*I; PII, *Pvu*II; S, *Sal*I. Invertase activity is expressed as units per 10^9 cells.

Inspection of the TUP1 protein-coding sequence revealed a pattern of conserved residues contained in about 43 amino acids (aa) which is repeated six times, comprising the carboxyl third of the protein (Fig. 5A). Certain residues are present in all or almost all repeats, allowing a consensus sequence to be deduced. This same pattern of conserved residues was previously reported in the repeated domains of the β subunit of transducin (transducin-β) and the yeast CDC4 protein (13, 56). The optimal alignment of four repeat regions of TUP1, CDC4, and transducin-β is shown in Fig. 5B. The consensus sequence among these three proteins is virtually identical to the consensus for the TUP1 repeats themselves. In the aligned regions depicted, TUP1 is 34% identical to CDC4 and 36% identical to transducin-β. If conservative amino acid substitutions are considered, in this region TUP1 and CDC4 are 56% identical or similar and TUP1 and transducin-β are 55.5% identical or similar. This high degree of sequence similarity indicates that TUP1 is evolutionarily related to these two proteins and that particular repeated domains of the proteins have been highly conserved.

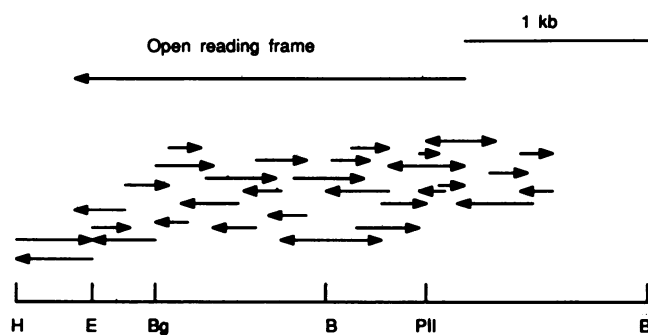**Construction and properties of deletion and insertion mu-**

FIG. 3. Sequencing strategy. Subclones in pBS or M13 were used for dideoxy sequencing of the *TUP1* region. Arrows denote size and directionality of each clone. Restriction sites: H, *Hind*III; E, *Eco*RI; Bg, *Bgl*II; B, *Bam*HI; PII, *Pvu*II.

tants. Deletion mutants of the *TUP1* gene were constructed by replacing a 3.1-kb *Eco*RI-*Bam*HI fragment or a 378-bp *Eco*RI-*Bgl*II fragment of the coding sequence with the *TRP1* gene (Fig. 6A). The larger deletion, *tup1-Δ1::TRP1*, removes the entire open reading frame and some upstream sequences. The smaller deletion, *tup1-Δ2::TRP1*, removes sequences downstream from the *Bgl*II site, including the region encoding the last 136 amino acids. An insertion mutation, *tup1-40::URA3*, was constructed by inserting the *URA3* gene at the *Eco*RI site near the 3' end of the coding sequence (Fig. 6A). This insertion introduced a stop codon into the coding sequence, thus truncating the encoded protein by 39 aa. RTY235 was transformed with these constructions to produce RTY418 (*tup1-Δ1::TRP1*), RTY420 (*tup1-Δ2::TRP1*), and RTY421 (*tup1-40::URA3*). The chromosomal structures of these strains were confirmed by Southern blots of genomic DNA extracted as described by Holm et al. (17; data not shown). All three mutant strains exhibited extreme *tup1* mutant phenotypes, including aggregation, self-shmooing, and total derepression of invertase activity (Fig. 6B). These results suggest that the C terminus of the TUP1 protein has a crucial functional role.

The *snf1* block on derepression of invertase is effectively bypassed in strains with point or deletion mutations in *cyc8*. It was subsequently found that a *tup1* point mutation was also capable of suppressing *snf1* (27). We tested the ability of a *tup1* deletion to suppress the *snf1* block of invertase derepression. Strain RTY469 (*snf1-Δ100::URA3*), bearing a deletion of the *SNF1* gene, was unable to derepress invertase, expressing less than 1 U of activity under both repressing and derepressing conditions. Isogenic strain RTY508 (*tup1-Δ1::TRP1 snf1-Δ100::URA3*), bearing deletions in both *SNF1* and *TUP1* genes, expressed 22 and 41 U invertase under repressing and derepressing conditions, respectively. Clearly, the *tup1* deletion completely suppresses the *snf1* effect on invertase expression.

The effect of *tup1* deletions on the mating efficiencies of *MAT*α cells was determined by the procedure of Lemontt et al. (21). Strains RTY418 (*tup1-Δ1::TRP1*), RTY420 (*tup1-Δ2::TRP1*), RTY421 (*tup1-40::URA3*), and RTY438 (*tup1-*

```
TGTACATCCCAAAGACATAAACATATGTATATAAGAACAACATAAACAAATAAATAAACCAGGAAAGCCCTAATTTACCCTTCTTTTTGGCTGAGC    96
AATCTAGGAGTTAATAATTTTTTGGAACTTTCCCCCTAACATTTTTAGAGACAGAACCCTCTTACATTATCGCTACCGACGGACCTTTAATTTTTC   192
GTTACTTCTCAAAAAACACTGAAAAACTGCGAAAAAATTAAAAAGCTCTCCCGTCAAAGCAACAAACAACAAAACAAAATAACAAGTACATAAGTA   288
AACAAATAGAGAGAAATTCCCCCATCTACCCCCAATTGATAAACATCAGCGAAGCAAGAACAACTGGCTGAACACGTGCCCCTCTATCTGTCCTTC   384
TGCAATTGTTTAATTTTCTTTCACGCCAAGTTACCTTTCGCACACCACTACTCTTTTTCTATTGTTTTTTTGTCTTTTTTGATAAGCAGGGGAAGA   480
AAGAAATCAGCTTTCCATCCAAACCAAT                                                                       509

ATG ACT GCC AGC GTT TCG AAT ACG CAG AAT AAG CTG AAT GAG CTT CTC GAT GCC ATC AGA CAG GAG TTT CTC   581
Met Thr Ala Ser Val Ser Asn Thr Gln Asn Lys Leu Asn Glu Leu Leu Asp Ala Ile Arg Gln Glu Phe Leu    24

CAA GTC TCA CAA GAG GCA AAT ACC TAC CGT CTT CAA AAC CAA AAG GAT TAC GAT TTC AAA ATG AAC CAG CAG   653
Gln Val Ser Gln Glu Ala Asn Thr Tyr Arg Leu Gln Asn Gln Lys Asp Tyr Asp Phe Lys Met Asn Gln Gln    48

CTG GCT GAG ATG CAG CAG ATA AGA AAC ACC GTC TAC GAA CTG GAG CTA ACT CAC AGG AAA ATG AAG GAC GCG   725
Leu Ala Glu Met Gln Gln Ile Arg Asn Thr Val Tyr Glu Leu Glu Leu Thr His Arg Lys Met Lys Asp Ala    72

TAC GAA GCA GAG ATC AAG CAC TTG AAA CTA GGG CTG GAG CAA AGA GAC CAT CAA ATT GCA TCT TTG ACC GTC   797
Tyr Glu Ala Glu Ile Lys His Leu Lys Leu Gly Leu Glu Gln Arg Asp His Gln Ile Ala Ser Leu Thr Val    96

CAG CAA CAG CAG CAA CAG CAA CAG CAG CAA CAG GTC CAG CAG CAT TTA CAA CAG CAA CAA CAG CAG CTA GCC   869
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Val Gln Gln His Leu Gln Gln Gln Gln Gln Gln Leu Ala   120

GCT GCA TCT GCA TCT GTT CCA GTT GCG CAA CAA CCA CCG GCT ACT ACT TCG GCC ACC GCC ACT CCA GCA GCA   941
Ala Ala Ser Ala Ser Val Pro Val Ala Gln Gln Pro Pro Ala Thr Thr Ser Ala Thr Ala Thr Pro Ala Ala   144

AAC ACA ACT ACT GGT TCG CCA TCG GCC TTC CCA GTA CAA GCT AGC CGT CCT AAT CTG GTT GGC TCA CAG TTG  1013
Asn Thr Thr Thr Gly Ser Pro Ser Ala Phe Pro Val Gln Ala Ser Arg Pro Asn Leu Val Gly Ser Gln Leu   168

CCT ACC ACC ACT TTG CCT GTG GTG TCC TCA AAC GCC CAA CAA CAA CTA CCA CAA CAG CAA CTG CAA CAG CAG  1085
Pro Thr Thr Thr Leu Pro Val Val Ser Ser Asn Ala Gln Gln Gln Leu Pro Gln Gln Gln Leu Gln Gln Gln   192

CAA CTT CAA CAA CAG CAA CCA CCT CCC CAG GTT TCC GTG GCA CCA TTG AGT AAC ACA GCC ATC AAC GGA TCT  1157
Gln Leu Gln Gln Gln Gln Pro Pro Pro Gln Val Ser Val Ala Pro Leu Ser Asn Thr Ala Ile Asn Gly Ser   216

CCT ACT TCT AAA GAG ACC ACT ACT TTA CCC TCT GTC AAG GCA CCT GAA TCT ACG TTG AAA GAA ACT GAA CCG  1229
Pro Thr Ser Lys Glu Thr Thr Thr Leu Pro Ser Val Lys Ala Pro Glu Ser Thr Leu Lys Glu Thr Glu Pro   240

GAA AAT AAT AAT ACC TCG AAG ATA AAT GAC ACC GGA TCC GCC ACC ACG GCC ACC ACT ACC ACC GCA ACT GAA  1301
Glu Asn Asn Asn Thr Ser Lys Ile Asn Asp Thr Gly Ser Ala Thr Thr Ala Thr Thr Thr Thr Ala Thr Glu   264

ACT GAA ATC AAA CCT AAG GAG GAA GAC GCC ACC CCG GCT AGT TTG CAC CAG GAT CAC TAC TTA GTC CCT TAT  1373
Thr Glu Ile Lys Pro Lys Glu Glu Asp Ala Thr Pro Ala Ser Leu His Gln Asp His Tyr Leu Val Pro Tyr   288

AAT CAA AGA GCA AAC CAC TCT AAA CCT ATC CCA CCT TTC CTT TTG GAT CTA GAT TCC CAG TCT GTT CCC GAT  1445
Asn Gln Arg Ala Asn His Ser Lys Pro Ile Pro Pro Phe Leu Leu Asp Leu Asp Ser Gln Ser Val Pro Asp   312

GCT CTG AAG AAG CAA ACA AAT GAT TAT TAT ATT TTA TAC AAC CCG GCA CTA CCA AGA GAA ATT GAC GTT GAG  1517
Ala Leu Lys Lys Gln Thr Asn Asp Tyr Tyr Ile Leu Tyr Asn Pro Ala Leu Pro Arg Glu Ile Asp Val Glu   336

TTA CAC AAA TCT TTG GAT CAT ACT TCA GTT GTT TGT TGC GTG AAG TTC AGT AAC GAT GGT GAA TAC TTA GCC  1589
Leu His Lys Ser Leu Asp His Thr Ser Val Val Cys Cys Val Lys Phe Ser Asn Asp Gly Glu Tyr Leu Ala   360

ACA GGC TGC AAC AAA ACT ACT CAA GTG TAT CGC GTT TCA GAC GGT TCT CTG GTG GCC CGT TTA TCT GAC GAT  1661
Thr Gly Cys Asn Lys Thr Thr Gln Val Tyr Arg Val Ser Asp Gly Ser Leu Val Ala Arg Leu Ser Asp Asp   384

TCT GCT GCC AAT AAC CAT CGA AAT TCG ATC ACT GAA AAT AAC ACC ACC ACG TCC ACG GAT AAC AAT ACA ATG  1733
Ser Ala Ala Asn Asn His Arg Asn Ser Ile Thr Glu Asn Asn Thr Thr Thr Ser Thr Asp Asn Asn Thr Met   408

ACA ACC ACT ACT ACC ACC ACA ATT ACT ACC ACA GCG ATG ACT TCG GCA GCA GAA TTG GCA AAA GAT GTG GAA  1805
Thr Thr Thr Thr Thr Thr Thr Ile Thr Thr Thr Ala Met Thr Ser Ala Ala Glu Leu Ala Lys Asp Val Glu   432

AAC CTG AAC ACT TCG TCT TCC CCA TCA TCC GAC TTG TAT ATC CGT TCA GTG TGT TTT TCT CCA GAT GGG AAA  1877
Asn Leu Asn Thr Ser Ser Ser Pro Ser Ser Asp Leu Tyr Ile Arg Ser Val Cys Phe Ser Pro Asp Gly Lys   456

TTT TTG GCA ACA GGT GCT GAA GAC AGA CTG ATT AGA ATT TGG GAT ATT GAA AAT AGA AAG ATT GTT ATG ATT  1949
Phe Leu Ala Thr Gly Ala Glu Asp Arg Leu Ile Arg Ile Trp Asp Ile Glu Asn Arg Lys Ile Val Met Ile   480

CTT CAA GGC CAC GAA CAA GAT ATT TAT TCA TTG GAC TAC TTT CCC TCA GGT GAC AAA TTA GTC TCC GGT TCT  2021
Leu Gln Gly His Glu Gln Asp Ile Tyr Ser Leu Asp Tyr Phe Pro Ser Gly Asp Lys Leu Val Ser Gly Ser   504

GGT GAC CGT ACC GTT CGT ATT TGG GAC TTA CGT ACA GGC CAG TGT TCA TTG ACT TTA TCC ATT GAA GAT GGT  2093
Gly Asp Arg Thr Val Arg Ile Trp Asp Leu Arg Thr Gly Gln Cys Ser Leu Thr Leu Ser Ile Glu Asp Gly   528

GTT ACC ACC GTC GCT GTA TCA CCA GGT GAT GGT AAA TAC ATC GCT GCT GGT TCT CTA GAT CGT GCT GTG AGA  2165
Val Thr Thr Val Ala Val Ser Pro Gly Asp Gly Lys Tyr Ile Ala Ala Gly Ser Leu Asp Arg Ala Val Arg   552

GTT TGG GAT TCC GAG ACC GGA TTC TTG GTG GAA AGA CTA GAT TCG GAA AAC GAA TCC GGT ACA GGC CAC AAG  2237
Val Trp Asp Ser Glu Thr Gly Phe Leu Val Glu Arg Leu Asp Ser Glu Asn Glu Ser Gly Thr Gly His Lys   576

GAC TCT GTT TAT AGC GTT GTC TTC ACG AGA GAT GGA CAA AGC GTT GTA TCC GGC TCA TTA GAT AGA TCT GTT  2309
Asp Ser Val Tyr Ser Val Val Phe Thr Arg Asp Gly Gln Ser Val Val Ser Gly Ser Leu Asp Arg Ser Val   600

AAG CTC TGG AAT TTG CAG AAT GCA AAC AAC AAG AGC GAT TCG AAA ACT CCA AAT TCC GGC ACT TGT GAA GTT  2381
Lys Leu Trp Asn Leu Gln Asn Ala Asn Asn Lys Ser Asp Ser Lys Thr Pro Asn Ser Gly Thr Cys Glu Val   624

ACG TAT ATC GGG CAT AAA GAC TTT GTA TTG TCC GTG GCC ACC ACA CAA AAT GAT GAG TAC ATC TTG TCC GGT  2453
Thr Tyr Ile Gly His Lys Asp Phe Val Leu Ser Val Ala Thr Thr Gln Asn Asp Glu Tyr Ile Leu Ser Gly   648

TCC AAA GAT CGT GGT GTC CTG TTT TGG GAT AAG AAA TCC GGC AAT CCG TTA TTG ATG TTG CAA GGT CAT AGG  2525
Ser Lys Asp Arg Gly Val Leu Phe Trp Asp Lys Lys Ser Gly Asn Pro Leu Leu Met Leu Gln Gly His Arg   672

AAT TCA GTT ATA TCT GTG GCT GTG GCA AAC GGG TCT TCG CTG GGT CCA GAA TAT AAC GTT TTT GCT ACT GGT  2597
Asn Ser Val Ile Ser Val Ala Val Ala Asn Gly Ser Ser Leu Gly Pro Glu Tyr Asn Val Phe Ala Thr Gly   696

AGC GGT GAT TGT AAA GCA AGG ATT TGG AAG TAT AAA AAA ATA GCG CCA AAT TAA                            2651
Ser Gly Asp Cys Lys Ala Arg Ile Trp Lys Tyr Lys Lys Ile Ala Pro Asn                                713

CAACAGAACACAAAAGGAACACTTTACAAATGTAACTAACTAAACTATTCTTCAATTCATTGTCATTATATCTTTCTCTCTATTCTTTTTTATCTA  2747
CTTATTTGTTATGTTATGATCCAGGTACGCTTTAAAACAAACAGTTACTTTCCTCACCATTCATCCCTTCCCCTCCTTTTCCTGTGTCTGAATATT  2843
GTGGTTCATTATTCTTTGATATGAATTTATTTGTGTATTGTATAAGCTT                                                   2892
```

FIG. 4. *TUP1* nucleotide sequence. The sequence of 2,902 nucleotides including *TUP1* is shown. Dots are placed over the start points of the *TUP1* mRNA, determined by primer extension at nucleotides 282, 284, 291, and 294. A putative transcription termination signal between nucleotides 2780 and 2790 is underlined. Polyglutamine and polythreonine domains are underlined.

6504

A

```
        S P S S D L Y I R S V C F - - - - - S D G K F L A T G A E D R L I R I W D  a (441-471)
        L Q G H E Q D L Y S L D Y F P - - - - S G D K L V S G S G D R T V R I W D  b (481-513)
        E D G V T T V A V S - - - - - - - - P G D G K Y I A A G S L D R A V R V W D  c (526-555)
        G T G H K D S V Y S V V - - - - - F T R G Q S V V S G S L D R S V K L W N  d (572-604)
        Y I G H K D F V L S V A - - - - T T Q N D E Y I L S G S K D R G V L F W D  e (626-655)
        L Q G H R N S V I S V A N G S S L G P E Y N V F A T G S G D C K L A R I W K  f (665-706)
```

B

```
TUP1           440 P S S D L Y I R S V C F S P D G K F L A T G A E D R L I R I W D 471
TRANSDUCIN-β   181 T G H T G D V M S L S L A P D T R L F V S G A C D A S A K L W D 212
CDC4           378 R G H M T S V I T C L Q F E D N Y V I - T G A D D K M I R V Y D 408

               482 Q G H E Q D I Y S L D Y F P S G D K L V S G S G D R T V R I W D 513
               223 T G H E S D I N A I C F F P N G N A F A T G S D D A T C R I F D 254
               419 S G H D G G V W A L K Y - A H G G I L V S G S T D R T V R V W D 449

               527 D G V T T - V A - V - S P G D G K Y I A A G S L D R A V R V W D 555
               271 C G L T S - V - S F - S K S - G R L L L A G Y D D F N C N V W D 298
               527 R G H M A S V R T V - S - G H G N I V V S G S Y D N T L I V W D 556

               573 T G H K D S V Y S V V F T R D G Q S V V S G S L D R S V K L W N 604
               309 A G H D N R V S C L G V T D D G M A V A T G S W D S F L K I W N 340
               567 S G H T D R L Y S T I Y D H E R K R C I S A S M D T T I R I W D 598
```

```
CONSENSUS        G H   V   S α         D G   α α S G S   D   α R α W D
                       I   T                         T           K
```

FIG. 5. Repeated domains of the TUP1 protein. The amino acid sequence of TUP1 shows a repeating structure of about 43 aa containing a pattern found in several other proteins. (A) The six repeats found at the C terminus of the predicted TUP1 protein (a to f); (B) comparison of four of the repeats in the predicted TUP1 protein with sequences of transducin-β and the product of *CDC4*. Numbers in parentheses denote the positions of amino acids within the predicted sequence. Amino acid residues within boxes conform to the consensus sequence. α, Aliphatic amino acids.

$\Delta 1::TRP1$ $cyc8-\Delta 1::LEU2$) had mating efficiencies of 2 × 10^{-4}, 1 × 10^{-4}, 2 × 10^{-5}, and 2.5 × 10^{-4}, respectively, compared with 0.6 for the isogenic wild-type strain RTY235. Thus, *tup1* deletions can induce sterility in *MAT*α strains, as was previously demonstrated for *tup1* point mutations.

To examine the functional relationship between *CYC8* and *TUP1*, isogenic strains with deletions in one or both genes were created by transformation of RTY235. The strains were grown under repressing conditions of 2% glucose, and their invertase activities were assayed. The single mutants RTY418 (*tup1-Δ1::TRP1*) and RTY490 (*cyc8-Δ1::LEU2*) expressed 59.6 and 67.6 U invertase activity, compared with 0.5 U for the wild-type strain RTY235. The double mutant RTY438 (*tup1-Δ1::TRP1 cyc8-Δ1::LEU2*) had 129 U, nearly twice as much activity as either of the single mutants. This difference suggests that some residual ability to repress transcription remains in the single mutants, implying that the individual gene products may have some activity in the absence of the other.

Effect of overexpression of *TUP1* and *CYC8* on invertase derepression. The wild-type strain RTY235 was transformed with multicopy plasmids carrying *CYC8*, *TUP1*, or both genes, and invertase activities of the untransformed and transformed strains grown under repressing and derepressing conditions were assayed. All of the strains had very low levels of invertase (<1 U per 10^9 cells) under repressing conditions, as expected. RTY235 expressed 40.6 U under

derepressing conditions. RTY235 transformed with pFW28 (*TUP1*) and RTY235 transformed with pRT81 (*CYC8*) expressed only 13.3 U (33% of the wild-type level) and 11.7 U (29% of the wild-type level), respectively, when cells were grown under derepressing conditions. The strain bearing plasmid pTXL63 (*CYC8* and *TUP1*) expressed only 7.6 U (18.7% of the wild-type level). These results extend those of the previous report (37) that invertase expression was reduced to 40% of derepressed levels by multiple copies of *CYC8*. These findings suggest that both *CYC8* and *TUP1* are directly involved in the process of glucose repression of transcription.

We have consistently observed that plasmids bearing the wild-type *CYC8* or *TUP1* gene do not completely complement *cyc8* or *tup1* mutations, using the criteria of constitutive invertase activity and flocculation. The ability of these genes on 2μm-based (multicopy) and centromere-based (single-copy) plasmids to complement deletion mutations was determined by invertase assays of cells grown under glucose-repressing conditions. When the *tup1* deletion strain RTY418 was transformed with the multicopy plasmid pFW28 bearing the *TUP1* gene, the invertase activity is reduced from 38 to 4.4 U per 10^9 cells, still greater than the wild-type repressed level of 0.6 (Table 2). A centromere-based plasmid carrying *TUP1*, pFW46, was even less effective in complementation. The complementation of *cyc8* presents a very similar pattern, with a multicopy plasmid

A

B

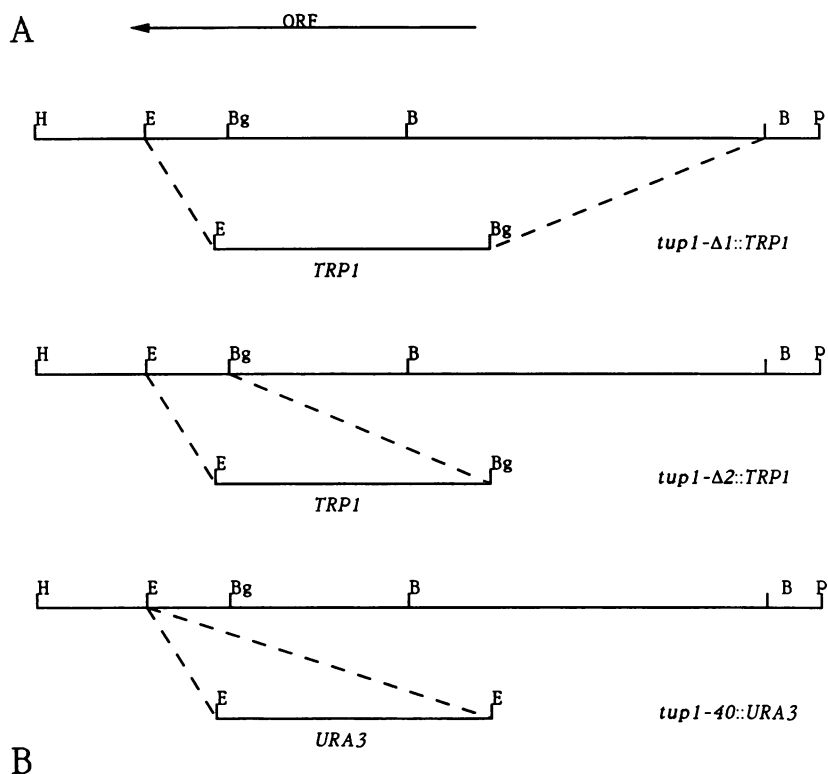| Genotype	Aggregation	Presence of shmoos	Invertase activity (units)	
wild-type	no	no	$<1/10^9$	cells
tup1-Δ1::TRP1	yes	yes	$42.3/10^9$	cells
tup1-Δ2::TRP1	yes	yes	$38.6/10^9$	cells
tup1-40::URA3	yes	yes	$34.6/10^9$	cells

FIG. 6. Deletion and insertion mutants. (A) Restriction fragments carrying the *TRP1* and *URA3* genes were used to interrupt the *TUP1* coding sequence at different sites. Either the whole gene (*tup1-Δ1::TRP1*), the coding sequence for 114 aa (*tup1-Δ2::TRP1*), or the coding sequence for 39 aa (*tup1-40::URA3*) was replaced. ORF, Open reading frame. (B) Aggregation, morphology, and invertase activity data for wild-type and mutant constructions. The cells used for the invertase assays were grown under repressing conditions (2% glucose).

providing better complementation than a single-copy plasmid, but neither restoring repression to the level of wild-type cells. The explanation for this partial complementation is unknown. It may be that a small percentage of cells that have lost the plasmids are derepressed and express invertase and cell surface components responsible for flocculation.

Identification of the *TUP1* mRNA. Northern blots of RNA from wild-type and mutant strains were used to identify the *TUP1* mRNA. The hybridization probe was a 126-bp *Bst*BI-

TABLE 2. Effects of copy number of *TUP1* and *CYC8* on complementation

Strain	Genotype	Plasmid	Invertase activity $(U/10^9 \text{ cells})$
RTY235	Wild-type	None	0.6
RTY363	Δcyc8	None	55.5
RTY363	Δcyc8	pFW46 (*CEN CYC8*)	11.9
RTY363	Δcyc8	pRT81 (2μm *CYC8*)	4.3
RTY418	Δtup1	None	38.0
RTY418	Δtup1	pFW45 (*CEN TUP1*)	11.7
RTY418	Δtup1	pFW28 (2μm *TUP1*)	4.4

*Pvu*II fragment from the protein-coding region, which was devoid of sequences encoding polyglutamine present in several other yeast genes (31, 39, 51). Northern blot analysis of total and poly(A)$^+$ RNA showed that the wild-type *TUP1* mRNA was polyadenylated and approximately 2.6 kb in length, as determined by comparison with labeled restriction fragments. Strain RTY234 (Fig. 7, lane 1) showed the size and level of expression of the *TUP1* mRNA in wild-type cells. A much stronger signal was seen in poly(A)$^+$ RNA (lane 2). Strain RTY359 (lane 3), a *tup1* mutant transformed with a plasmid containing *TUP1*, showed the wild-type size of the *TUP1* mRNA. Strain RTY363, a *cyc8* deletion strain, also expressed *TUP1* mRNA at a level similar to the wild-type level (lane 4). This strain was examined because *cyc8* and *tup1* mutants have very similar phenotypes, and it was conceivable that one gene could affect the expression of the other. Strains RTY336 and RTY338 are transposon insertional mutants with extreme mutant phenotypes (Fig. 2). Both strains yielded truncated mRNAs and had lower levels of expression of *TUP1* mRNA (lanes 5 and 6). Strain RTY418 is a total *tup1* deletion strain with the *TRP1* gene inserted at the *TUP1* locus. This strain had no detectable
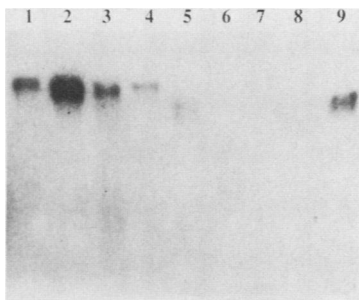
FIG. 7. Northern blot of wild-type and mutant *TUP1* mRNAs. RNA purified from wild-type and *tup1* mutant strains was electrophoresed on a denaturing 0.8% agarose gel and transferred to nitrocellulose (see Materials and Methods). All lanes contained 10 μg of total cellular RNA except lane 2, which was loaded with 3 μg of poly(A)$^+$ RNA. The RNA was transferred to nitrocellulose and probed with a *Pvu*II-*Bst*BI fragment (nucleotides 523 to 649) from the *TUP1* coding region, which was labeled with ^{32}P by nick translation. Lanes: 1, RTY234 (wild type); 2, RTY234, 3 μg of poly(A)$^+$ RNA; 3, RTY110 (*tup1-100*, transformed with pFW28); 4, RTY363 (*cyc8Δ-1*); 5, RTY336 (*tup1-11::URA3*); 6, RTY338 (*tup1-7::URA3*); 7, RTY418 (*tup1Δ-1*); 8, RTY420 (*tup1Δ-2*); 9, RTY421 (*tup1-40::URA3*).

TUP1 mRNA, confirming that it is a null mutant (lane 7). Strain RTY420 is a C-terminal deletion that replaces a 378-bp *Eco*RI-*Bgl*II fragment with the *TRP1* gene. This strain made very little stable *TUP1* message. The message was also clearly truncated (lane 8). RTY421, which contains an insertion of the *URA3* gene into the *Eco*RI site near the end of the *TUP1* open reading frame, produced a truncated transcript in amount similar to that of the wild type (lane 9).

Mapping of the 5' end of the *TUP1* mRNA. The 5' ends of the *TUP1* mRNA were determined by primer extension and by RNase protection. For primer extension, a 17-base oligonucleotide corresponding to the beginning of the open reading frame was synthesized. This primer was hybridized to poly(A)$^+$ RNA from yeast cells overproducing the *TUP1* mRNA and extended with reverse transcriptase (Fig. 8A). The extension products were heterogeneous, with the strongest bands at positions 282, 284, 291, and 294, determined by comparison with a DNA sequencing reaction using the same primer and single-stranded bacteriophage DNA containing this region of *TUP1*. Identical results were obtained with two independent preparations of RNA (lanes 1 and 3), indicating that RNA degradation is an unlikely explanation for the heterogeneous ends. No bands were seen with use of RNA from a strain deleted for the *TUP1* gene (lane 2).

The 5' ends of the *TUP1* mRNA were also determined by RNase protection. A ^{32}P-labeled RNA probe containing bases 22 to 747 of the *TUP1* sequence was hybridized to total cellular RNA. The hybridized RNA was treated with RNases and separated by electrophoresis on a sequencing gel. In wild-type cells (Fig. 8B, lanes 1 and 2), multiple bands were found in the same region seen in the primer extension experiment (Fig. 8A). The position of the strongest band estimated by comparison with the RNA markers was 279, compared with 281 determined by primer extension. In addition, there was a smaller strong band estimated to be at position 489, 20 bases upstream of the beginning of the open reading frame at 509. No major start sites were found in this region by primer extension (data not shown). The same bands were found in RNA of a strain bearing *TUP1* on a multicopy plasmid, but they were amplified severalfold (Fig. 8B, lanes 5 and 6). When the RNase concentrations were

increased twofold (even-numbered lanes), the same pattern of bands was observed, with diminished intensities. When RNA from a *tup1* deletion strain was used, no bands were seen (lanes 3 and 4). As a control, the RNase protection mapping of the 5' end of the *CYC8* mRNA that was reported previously (51) was performed at the same time, using identical conditions and RNA samples (lanes 7 to 12). The previous results were confirmed, demonstrating that the multiple start sites for the *TUP1* mRNA are not due to RNA degradation or other artifacts.

DISCUSSION

This report describes the characterization of the yeast *TUP1* gene, which plays a major role in the glucose repression of transcription. In addition, *tup1* mutants display a number of phenotypes, including flocculation, sporulation deficiency, and *MAT*α sterility. The mechanism whereby *TUP1* and the functionally related *CYC8* gene regulate these multiple pathways has been unclear. A plausible hypothesis is that they inhibit a class of transcriptional activator proteins.

The major 5' ends of the *TUP1* mRNA were mapped to a cluster of sites in the region of nucleotides 280 to 290 and to a site near nucleotide 489. The large number of different 5' termini of the *TUP1* mRNA is unusual but not unprecedented in *S. cerevisiae*. A similar degree of heterogeneity was found for the 5' ends of the *CYC1* (14), *TRP2* (57), and *TRP3* (57) mRNAs. There are no sequences matching the consensus TATA box (TATAAA) upstream of these sites, but potential variant TATA boxes found at 180 to 187 (CCTTTAAT) and at 391 to 398 (TGTTTAAT) resemble a new class of TATA boxes with the consensus (C/T)ATT TAAt recently recognized by Singer et al. (42). The distance between the TATA box and the start point of transcription is usually between 40 and 120 bp in *S. cerevisiae* (45). A sequence conforming to the consensus for transcription termination of Zaret and Sherman (58) is found between nucleotides 2757 and 2780. The putative polyadenylation site would be downstream of this signal, giving an mRNA with a size of 2.5 to 2.6 kb, agreeing with the independently determined value of 2.6 kb from Northern blot analysis (Fig. 7, lane 1). All of the major start sites of the *TUP1* mRNA were mapped to sites upstream of the AUG at the beginning of the open reading frame at position 510. When a truncated *TUP1* open reading frame starting from the AUG at position 663 was expressed from the *ADH1* promoter, the gene product detected by anti-TUP1 antibodies was significantly truncated, demonstrating that the first AUG in the *TUP1* open reading frame is normally used (unpublished data).

A strain carrying a deletion of the entire *TUP1* open reading frame and upstream sequences made no *TUP1* mRNA when probed on a Northern blot (Fig. 7, lane 7) and is considered a null mutant. This mutant is viable and exhibits more severe versions of the same mutant phenotypes observed in point mutants. The deletion mutant exhibited extreme flocculence, higher invertase activities under glucose repressing conditions, and *MAT*α shmooing.

Two deletions made at the 3' end of the open reading frame (Fig. 6) induced phenotypes similar to those of the null mutant. These two mutations truncated the open reading frame by 114 and 39 aa. The latter mutation, *tup1-40::URA3*, removes the coding region for one of the six repeats having strong homology to the β subunit of G proteins. A strain bearing the *tup1-40::URA3* mutation made amounts of a truncated mRNA similar to wild-type levels (Fig. 7, lane 9).
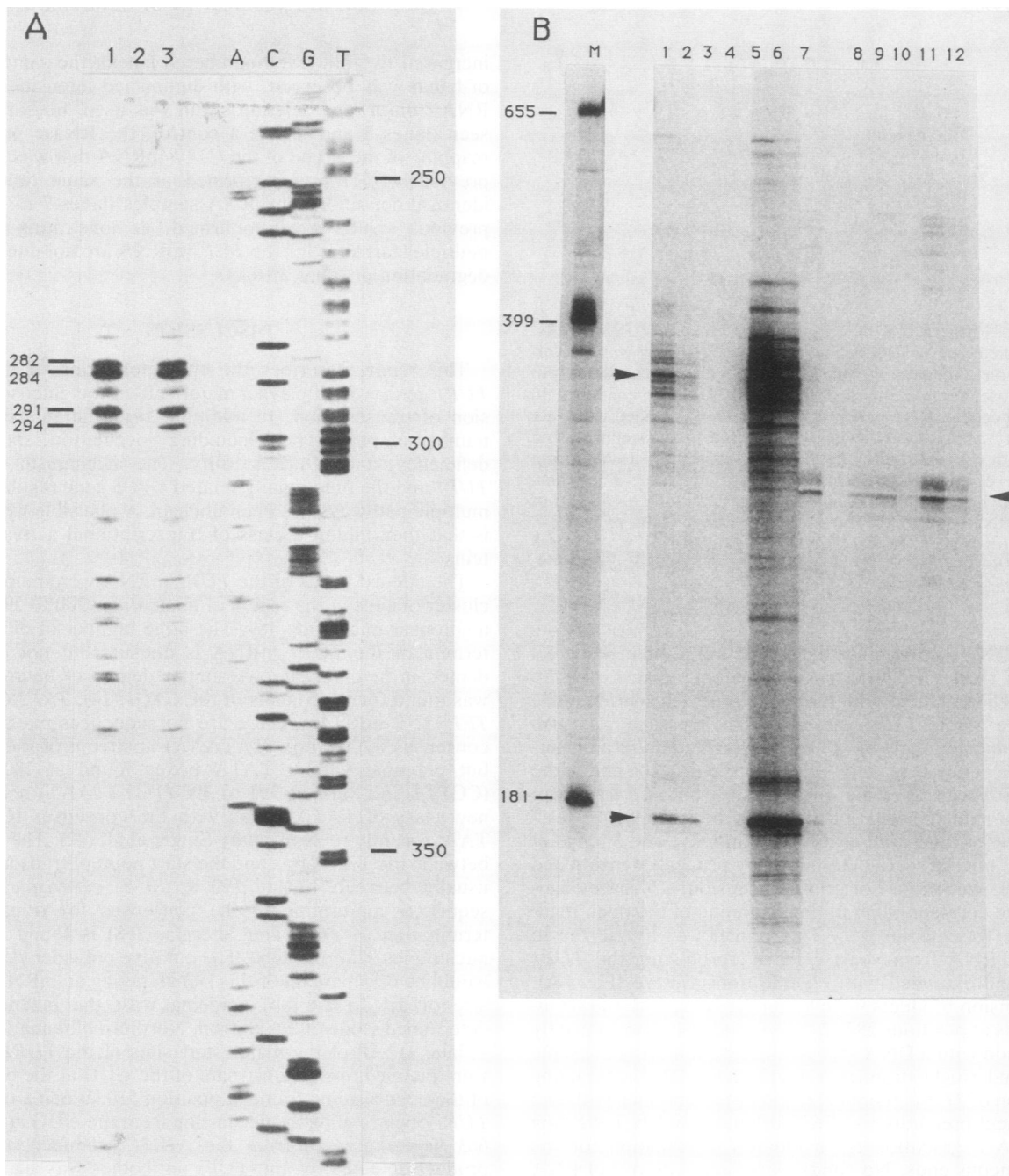
FIG. 8. Mapping of the 5' terminus of the *TUP1* mRNA. (A) Primer extension. A 17-base primer, 5'-CGAAACGCTGGCAGTCA-3', complementary to the mRNA between nucleotides 510 and 526 of the *TUP1* sequence was labeled at the 5' end with $[\gamma\text{-}^{32}P]ATP$. Hybridization of the primer with poly(A)$^+$ RNA and primer extensions were carried out as described in Materials and Methods. As a control for RNA degradation, two preparations of RNA purified from yeast strain RTY448 at different times were used (lanes 1 and 3). The primer extension reactions were separated on a standard 6% acrylamide sequencing gel. As a size standard, DNA sequencing reactions were carried out by using the same labeled primer and a single-stranded DNA template consisting of a 1.1-kb restriction fragment extending from a *Bst*BI site at nucleotide 523 of the *TUP1* sequence to the upstream *Bam*HI site subcloned into the *Acc*I-*Bam*HI sites of M13mp18. The sequence positions of the four strongest primer extension products and of the DNA sequencing reactions are shown. Lanes: 1, RNA from RTY448 (RTY418 transformed with pFW28, *TUP1* on YEp24), preparation 1; 2, RNA from RTY418 (Δ*tup1*); 3, RNA from RTY448, preparation 2; A, C, G, and T, DNA sequencing reactions. (B) RNase protection. Mapping of the 5' ends of the *TUP1* mRNA was performed as described previously (51). The template for RNA probe synthesis, pRT147, consisted of a 1.2-kb fragment from the *Pvu*II site at position 649 of the *TUP1* sequence to the *Bam*HI site upstream, subcloned into the *Hinc*II-*Bam*HI sites of pBS(+). The template DNA, pRT147 digested with *Nde*I (nucleotide 22 of the *TUP1* sequence), was transcribed with bacteriophage T3 RNA polymerase to produce a ^{32}P-labeled antisense transcript containing bases 22 to 747 of the *TUP1* sequence. Hybridization reaction mixtures contained 25 μg of total cellular RNA and 5 × 10^5 cpm of probe RNA. Standard RNase digestions (lanes 1, 3, and 5) used 40 μg of RNase A and 2 μg of RNase T_1 per ml; 2× digestions (lanes 2, 4, and 6) contained twice this amount. Control reactions using an RNA probe for the 5' end of the *CYC8* mRNA (51; Fig. 5A) were performed in parallel, using the same RNA samples and experimental conditions (lanes 7 to 12). Lanes: M, RNA size markers produced by transcription of pRT147; 1 and 2, RNA from X2180-1A (wild type); 3 and 4, RNA from RTY418 (Δ*tup1*); 5 and 6, RNA from RTY448 (RTY418 transformed with pFW28 (*TUP1* in YEp24); 7 to 12, same RNA samples and conditions as for 1 to 6, using the *CYC8* probe.

Since the mRNA has the normal 5' sequences and should be efficiently translated, normal amounts of the truncated protein are probably made, supporting the idea that the C terminus is required for *TUP1* function. Confirmation of this conclusion will require evidence that the truncated proteins are actually synthesized.

To investigate the relationship between the two genes, strains bearing deletions in the *CYC8* or *TUP1* gene were compared. Their phenotypes appeared to be identical for all of the characters examined: high-level constitutive invertase activity, lack of sporulation, *MAT*α sterility and shmooing, flocculation, and ability to suppress the *snf1* block on invertase derepression. A strain with deletions in both the *CYC8* and *TUP1* genes had properties very similar to those of the single mutants, with somewhat higher invertase activity.

The predicted TUP1 protein is rich in glutamine, serine, and threonine. Polyglutamine and polythreonine tracts are found in the 5' end and middle of the coding region, respectively. Serine and threonine residues make up approximately 20% of the predicted protein and are potential sites for covalent modifications such as phosphorylation. Polyglutamine sequences have been found in a rapidly increasing number of proteins, almost all of them regulatory proteins. Most of these proteins are believed to bind to DNA because of the presence of well-defined DNA-binding motifs. Included in this class are the yeast regulatory proteins HAP2 (31), PHO2 (39), GAL11 (47), ADR6 (29), and MCM1 (30) and many of the *Drosophila* proteins regulating embryonic development (32, 52). An exception is the *Drosophila* Notch protein, which regulates neural development but appears to be located in the plasma membrane (52). Although the function of polyglutamine is not known, it would probably form an α helix (8) and might possibly connect domains of a protein by acting as a flexible arm or tether. Polythreonine sequences have been found in the *Drosophila* Zipper protein and in the Sgs-3 glue protein (24, 59).

The carboxyl third of the predicted TUP1 protein is composed of six repeats of about 43 aa each, with specific residues highly conserved among the repeats. In a search of the protein data base, this region matched similar repeated domains in transducin-β and the yeast CDC4 protein. The CDC4 protein is required for progression beyond the G_1 phase of the cell cycle, but its precise functional role is unknown. The consensus sequence for the TUP1 repeats (Fig. 5A) is virtually identical to the consensus for the alignment between TUP1, CDC4, and transducin-β and to the consensus previously reported for the CDC4 and transducin-β repeats (13). Two recent additions to this family are the yeast STE4 protein (53), which is involved in the cellular response to mating pheromones, and the protein encoded by the *Drosophila* gene Enhancer of split (15), implicated in neurogenesis. All of these related proteins have six copies of the repeated domains, which are always found at the carboxyl terminus. The proteins fall into two distinct size classes. The yeast STE4 protein and transducin-β are 423 and 340 aa in length, respectively, and are composed almost entirely of the repeat domains. Both proteins serve as β subunits in the heterotrimeric G-protein complexes which relay signals from hormone receptors (or rhodopsin, in the case of transducin) to effector proteins such as adenylyl cyclase. In mammals there are two forms of the β subunit, β_1 and β_2, encoded by separate genes. Only the β_1 form is found in transducin, while both the β_1 and β_2 forms are found in other G proteins. The properties of the β_1 and β_2 forms are very similar, and it appears that most of the

functional differences among various G proteins are due to heterogeneity in the α and γ subunits (26). The larger proteins, including TUP1 (713 aa), CDC4 (779 aa), and Enhancer of split (719 aa), are not known to form part of a G-protein complex, and their biochemical roles are still undefined. We have not found any significant homology between these pr' teins outside of the repeated domains.

Two major roles for the β subunit in G-protein function have been postulated. According to the most widely accepted model, the βγ subunits play a negative role by inhibiting or attenuating the activity of the α subunit (26). An activated receptor interacts with the G protein, resulting in an exchange of GDP for GTP bound to the α subunit. The α subunit then disassociates from the βγ subunits and modulates the activity of an effector protein (46). However, in some cases, such as the opening of atrial potassium channels (22), the βγ subunits may play a positive role by directly interacting with the effector. Very strong evidence for a positive role for the βγ subunits comes from the genetic interaction of mutations affecting the yeast G-protein subunits involved in the mating pheromone response pathway (53). Mutation in *STE4*, encoding the β subunit, or *STE18*, encoding the γ subunit, can suppress the cell cycle arrest phenotype of mutations in *SCG1*, which encodes the α subunit. These results imply that the β and γ subunits transmit the signal for cell cycle arrest, and the role of the α subunit is to inhibit or regulate this activity.

The homology of the TUP1 protein to the β subunit of G proteins suggests that TUP1 acts in a signal transduction pathway, rather than playing a passive or structural role. By analogy with the role of G proteins, we would expect the TUP1 repeat domains to be involved in protein-protein interactions. The TUP1 protein could inhibit a pathway required for activation of transcription of glucose-repressible genes, or it could directly inhibit transcription by binding to transcription factors. Possible targets for inhibition by TUP1 are the products of the *SNF2*, *SNF5*, and *SNF6* genes, which are required for the high-level expression of glucose-repressible genes (27).

Other roles for TUP1 and CYC8 in glucose repression have been considered. The possibility that they repress transcription by binding to DNA has not been excluded by experimental evidence, but none of the increasing number of motifs typical of DNA-binding proteins has been recognized in either the CYC8 or TUP1 protein sequence. Since mutations in either *CYC8* or *TUP1* suppress *snf1* mutations, it was conceivable that they could have a protein phosphatase activity that would counteract the SNF1 kinase activity. However, no similarity between the TUP1 or CYC8 protein and the recently published protein phosphatase sequences has been found.

The *tup1* deletion mutants described in this study had phenotypes identical to those of *cyc8* null mutants described previously (37, 51), reinforcing the conclusion that these two genes perform a single function. It seems likely that this function is the inhibition of certain transcriptional activators or transcription factors. The presence on the TUP1 protein of repeated domains characteristic of the β subunits of G proteins suggests a role for protein-protein interactions in TUP1 function. The factors that regulate TUP1 and CYC8 activities and the mechanisms whereby they repress transcription remain to be elucidated.

for isolation of the *SNF1* gene and Xianqiang Li for construction of plasmids pTXL6 and pTXL63.

LITERATURE CITED

1. **Akhurst, R. J., F. J. Calzone, J. J. Lee, R. J. Britten, and E. H. Davidson.** 1987. Structure and organization of the CyIII actin gene subfamily of the sea urchin *Strongylocentratus purpuratus.* J. Mol. Biol. **194:**193–203.
2. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligo-thymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA **69:**1408–1412.
3. **Bennetzen, J. L., and B. D. Hall.** 1982. Codon selection in yeast. J. Biol. Chem. **257:**3026–3031.
4. **Botstein, D., S. C. Falco, S. E. Stewart, M. Brennan, S. Scherer, D. T. Stinchcomb, K. Struhl, and R. W. Davis.** 1979. Sterile host yeasts (SHY): a eukaryotic system of biological containment for recombinant DNA experiments. Gene **8:**17–24.
5. **Carlson, M., B. C. Osmond, L. Neigeborn, and D. Botstein.** 1984. A suppressor of *snf1* mutations causes constitutive high-level invertase synthesis in yeast. Genetics **107:**19–32.
6. **Celenza, J. L., and M. Carlson.** 1986. A yeast gene that is essential for release for glucose repression encodes a protein kinase. Science **233:**1175–1180.
7. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4:**165–170.
8. **Chou, P. Y., and G. D. Fasman.** 1974. Prediction of protein conformation. Biochemistry **13:**222–245.
9. **Ciriacy, M.** 1977. Isolation and characterization of yeast mutants defective in intermediary carbon metabolism and in carbon catabolite derepression. Mol. Gen. Genet. **154:**213–220.
10. **Coerces, V., A. Pellicer, R. Axel, and M. Meselson.** 1981. Integration, transcription, and control of a *Drosophila* heat shock gene in mouse cells. Proc. Natl. Acad. Sci. USA **78:**7038–7042.
11. **Denis, C. L.** 1984. Identification of new genes involved in the regulation of yeast alcohol dehydrogenase II. Genetics **208:**101–106.
12. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:**6–13.
13. **Fong, H. K. W., J. B. Hurley, R. S. Hopkins, R. Miake-Lye, M. S. Johnson, R. F. Doolittle, and M. I. Simon.** 1986. Repetitive segmental structure of the transducin β subunit: homology with the *CDC4* gene and identification of related mRNAs. Proc. Natl. Acad. Sci. USA **83:**2162–2166.
14. **Guarente, L., and E. Hoar.** 1984. Upstream activation sites of the *CYC1* gene of *Saccharomyces cerevisiae* are active when inverted but not when placed downstream of the "TATA box." Proc. Natl. Acad. Sci. USA **81:**7860–7864.
15. **Hartley, D. A., A. Preiss, and S. Artavanis-Tsakonas.** 1988. A deduced gene product from the *Drosophila* neurogenic locus, *Enhancer of split*, shows homology to mammalian G-protein β subunit. Cell **55:**785–795.
16. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:**351–359.
17. **Holm, C., D. W. Meeks-Wagner, W. L. Fangman, and D. Botstein.** 1986. A rapid, efficient method for isolating DNA from yeast. Gene **42:**169–173.
18. **Ito, H., Y. Fakuda, K. Murata, and A. Kimura.** 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **153:**163–168.
19. **Kay, R., and J. McPherson.** 1987. Hybrid pUC vectors for addition of new restriction enzyme sites to the ends of DNA fragments. Nucleic Acids Res. **15:**2778.
20. **Lemontt, J. F.** 1977. Pathways of ultraviolet mutability in *Saccharomyces cerevisiae*. III. Genetic analysis and properties

of mutants resistant to ultraviolet induced forward mutations. Mutat. Res. **43:**179–203.
21. **Lemontt, J. F., D. R. Fugit, and V. L. MacKay.** 1980. Pleiotropic mutations at the *TUP1* locus that affect the expression of mating-type dependent functions in *Saccharomyces cerevisiae*. Genetics **94:**899–920.
22. **Logothetis, D. E., Y. Kurachi, J. Galper, E. J. Neer, and D. E. Clapham.** 1987. The βγ subunits of GTP-binding proteins activate the muscarinic K⁺ channel in heart. Nature (London) **325:**321–326.
23. **MacKay, V. L.** 1983. Cloning of yeast *STE* genes in 2μm vectors. Methods Enzymol. **101:**325–343.
24. **Martin, C. H., C. A. Mazeda, and E. M. Meyerowitz.** 1988. Evolution and expression of the Sgs-3 glue gene of Drosophila. J. Mol. Biol. **201:**273–287.
25. **Miesfield, R., S. Rusconi, P. J. Godowski, B. A. Maler, S. Okret, A. C. Wikstrom, J. A. Gustafson, and K. R. Yamamoto.** 1986. Genetic complementation of a glucocorticoid receptor deficiency by expression of cloned receptor cDNA. Cell **46:**389–399.
26. **Neer, E. J., and D. E. Clapham.** 1988. Roles of G protein subunits in transmembrane signalling. Nature (London) **333:**129–134.
27. **Neigeborn, L., and M. Carlson.** 1987. Mutations causing constitutive invertase synthesis in yeast: genetic interactions with *snf* mutations. Genetics **115:**247–253.
28. **Oberbaumer, I.** 1986. New pUC-derived expression vectors for rapid construction of cDNA libraries. Gene **49:**81–91.
29. **O'Hara, P. J., H. Horowitz, G. Eichinger, and E. T. Young.** 1988. The yeast *ADR6* gene encodes homopolymeric amino acid sequences and a potential metal binding domain. Nucleic Acids Res. **16:**10153–10169.
30. **Passemore, S., G. T. Maine, R. Elble, C. Christ, and B. K. Tye.** 1988. *Saccharomyces cerevisiae* protein involved in plasmid maintenance is necessary for mating of MAT alpha cells. J. Mol. Biol. **204:**593–606.
31. **Pinkham, J. L., J. T. Olesen, and L. P. Guarante.** 1987. Sequence and localization of *Saccharomyces cerevisiae* HAP2 protein, a transcriptional activator. Mol. Cell. Biol. **7:**578–585.
32. **Poole, S. J., L. M. Kauvar, B. Drees, and T. Kornberg.** 1985. The *engrailed* locus of Drosophila: structure and analysis of an embryonic transcript. Cell **40:**37–43.
33. **Rodgers, S., R. Wells, and M. Rechsteiner.** 1986. Amino acid sequences common to rapidly degraded protein: the PEST hypothesis. Science **234:**364–368.
34. **Rothstein, R. J., and F. Sherman.** 1980. Genes affecting the expression of cytochrome c in yeast: genetic mapping and genetic interactions. Genetics **94:**871–889.
35. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
36. **Schamhart, D. H. J., A. M. A. TenBerge, and K. W. VandePoll.** 1975. Isolation of a catabolite repression mutant of yeast as a revertant of a strain that is maltose negative in a respiratory deficient state. J. Bacteriol. **121:**747–752.
37. **Schultz, J., and M. Carlson.** 1987. Molecular analysis of *SSN6*, a gene functionally related to the *SNF1* protein kinase of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7:**3637–3645.
38. **Seiffert, H. L., E. Y. Chen, M. So, and F. Heffron.** 1986. Shuttle mutagenesis: a method of transposon mutagenesis for *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **83:**735–739.
39. **Sengstag, C., and A. Hinnen.** 1987. The sequence of the *Saccharomyces cerevisiae* gene *PHO2* codes for a regulatory protein with an unusual amino acid composition. Nucleic Acids Res. **15:**233–246.
40. **Sherman, F., G. R. Fink, and J. B. Hicks.** 1982. Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. **Sikorski, R. S., and P. Heiter.** 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. Genetics **122:**19–27.
42. **Singer, V. L., C. R. Wobbe, and K. Struhl.** 1990. A wide variety of DNA sequences can functionally replace a yeast TATA

element for transcriptional activation. Genes Dev. **4:**636–645.

43. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.

44. **Stark, H. C., D. Fugit, and D. B. Mowshowitz.** 1980. Pleiotropic properties of a yeast mutant insensitive to catabolite repression. Genetics **94:**921–928.

45. **Struhl, K.** 1987. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. Cell **49:**295–297.

46. **Stryer, L.** 1986. Cyclic GMP cascade of vision. Annu. Rev. Neurosci. **9:**87–119.

47. **Suzuki, Y., Y. Nogi, A. Abe, and T. Fukasawa.** 1988. *GAL11* protein, an auxiliary transcription activator for genes encoding galactose-metabolizing enzymes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8:**4991–4999.

48. **Tabor, S., and C. C. Richardson.** 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84:**4767–4771.

49. **Thrash-Bingham, C., and W. L. Fangman.** 1989. A yeast mutation that stabilizes a plasmid bearing a mutated *ARS1* element. Mol. Cell. Biol. **9:**809–816.

50. **Trumbly, R. J.** 1986. Isolation of *Saccharomyces cerevisiae* mutants constitutive for invertase synthesis. J. Bacteriol. **166:**1123–1127.

51. **Trumbly, R. J.** 1988. Cloning and characterization of the *CYC8* gene mediating glucose repression in yeast. Gene **73:**97–111.

52. **Wharton, K. A., K. M. Johansen, T. Xu, and S. Artavanis-Tsakonas.** 1985. Nucleotide sequence from the neurogenic locus *Notch* implies a gene product that shares homology with proteins containing EGF-like repeats. Cell **43:**567–581.

53. **Whiteway, M., L. Hougan, D. Dignard, D. Y. Thomas, L. Bell, G. C. Saari, F. J. Grant, P. O'Hara, and V. L. MacKay.** 1989. The *STE4* and *STE18* genes of yeast encode potential β and γ subunits of the mating factor receptor-coupled G protein. Cell **56:**467–477.

54. **Wickner, R. B.** 1974. Mutants of *Saccharomyces cerevisiae* that incorporate deoxythymidine-5'-monophosphate in deoxyribonucleic acid in vivo. J. Bacteriol. **117:**252–260.

55. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.

56. **Yochem, J., and B. Byers.** 1987. Structural comparison of the yeast cell division cycle gene *CDC4* and a related pseudogene. J. Mol. Biol. **195:**233–245.

57. **Zalkin, H., J. L. Paluh, M. van Cleemput, W. S. Moye, and C. Yanofsky.** 1987. Nucleotide sequence of *Saccharomyces cerevisiae* genes *TRP2* and *TRP3* encoding bifunctional anthranilate synthase:indole-3-glycerol phosphate synthase. J. Biol. Chem. **259:**3985–3992.

58. **Zaret, K. S., and F. Sherman.** 1982. DNA sequence required for efficient transcription termination in yeast. Cell **28:**563–573.

59. **Zhao, D. B., S. Cote, F. Jahnig, J. Haller, and H. Jackle.** 1988. *Zipper* encodes a putative integral membrane protein required for normal axon patterning during Drosophila neurogenesis. EMBO J. **7:**1115–1119.