MOLECULAR AND CELLULAR BIOLOGY, Sept. 1990, p. 4935–4941
0270-7306/90/094935-07$02.00/0

Vol. 10, No. 9

SPT6, an Essential Gene That Affects Transcription in *Saccharomyces cerevisiae*, Encodes a Nuclear Protein with an Extremely Acidic Amino Terminus

MICHELE S. SWANSON,[1,2] MARIAN CARLSON,[2] AND FRED WINSTON[1]*

Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115,[1] and Department of Genetics and Development and Institute of Cancer Research, Columbia University, College of Physicians and Surgeons, New York, New York 10032[2]

Received 5 March 1990/Accepted 31 May 1990

***SPT6* is an essential gene of *Saccharomyces cerevisiae* that appears to play a role in transcription. Mutations in the *SPT6* (*SSN20*, *CRE2*) gene suppress δ insertion mutations in the 5′ regions of *HIS4* and *LYS2* and mutations in *cis*- and/or *trans*-acting elements that are required for expression of *SUC2* and *ADH2*. We report here that *SPT6* encodes a 170-kilodalton highly charged protein with an extremely acidic amino terminus. By use of an epitope-tagged SPT6 protein, we have determined by indirect immunofluorescence that the SPT6 protein is located in the nucleus.**

Mutations in the *Saccharomyces cerevisiae* gene *SPT6* (also called *SSN20* [25, 26] and *CRE2* [7, 8]) have been isolated as suppressors of both *cis*- and *trans*-acting mutations that alter transcription of particular genes. Mutations in *SPT6* were first isolated as suppressors of an insertion mutation at the *HIS4* gene (*his4-912δ*) caused by a long terminal repeat of the transposable element Ty (32). Dominant and recessive mutations in *SPT6* (called *SSN20*) were also isolated as extragenic suppressors of the defect in *SUC2* expression caused by *snf2* or *snf5* mutations and were shown to act as suppressors of *snf6* (26) and *SUC2* mutant alleles with upstream activation sequence deletions (25). One *spt6* mutation, called *cre2-1*, was isolated in a selection for mutations that allow *ADH2* to escape glucose repression. The *cre2-1* mutation also restores a moderate level of *ADH2* expression in the absence of the *trans*-acting positive activator *ADR1* (7, 8). In every case that has been examined, suppression by *spt6* mutations is mediated at the level of transcription (6, 25).

The *SPT6* gene was cloned by complementation of both an *spt6* and an *ssn20* allele, and a 4.6-kilobase (kb) mRNA was identified (6, 25). The gene was mapped to the right arm of chromosome VII (6, 25). Null alleles of *SPT6* were constructed and found to be lethal in haploid cells (6, 25), consistent with the temperature-sensitive lethality conferred by many *spt6* alleles (7, 26, 32). Experiments with the cloned *SPT6* gene demonstrated that either increased or decreased dosage of the wild-type *SPT6* gene also suppresses *his4-912δ* mutations (6) and *SUC2* upstream regulatory region deletions (25). These results suggested that the SPT6 protein functions as part of a complex which is affected by altered stoichiometry (6, 25). Since *SPT6* appears to play an essential and general role in transcription, we have characterized further the gene and its product.

Sequence of the *SPT6* gene. The nucleotide sequence of the *SPT6* gene was determined by the method of Sanger et al. (30). The open reading frame (Fig. 1) extended for 4353 base pairs, a size consistent with the 4.6-kb *SPT6* RNA identified previously (6, 25). The nucleotide sequence predicted a highly charged 1,451-amino-acid protein of 168,290 daltons (Da). In the entire protein, 19% of the residues were acidic and 15% were basic. A striking characteristic of the predicted protein was its acidic amino terminus: the first 70 residues had a net charge of −30 (50% acidic residues), and the amino-terminal third of the protein (484 residues) had a net charge of −81 (29% acidic residues). In addition, the SPT6 sequence included several consensus sites for phosphorylation by casein kinase II (20; reviewed in reference 2; G. Prelich and F. Winston, unpublished observations).

Acidic residues are characteristic of the activation domains of the yeast transcriptional activators GAL4 (21) and GCN4 (16, 17). However, the negatively charged region of SPT6 is more extensive and more acidic than those of previously characterized transcription activation domains. Acidic regions have also been found in several proteins that are thought to interact with chromatin (see reference 10 for review).

Data base searches (National Biomedical Research Foundation Protein Sequence Database release 21 and translated GenBank version 62) did not reveal proteins with significant sequence similarity to SPT6.

Identification of the SPT6 protein. To identify the SPT6 protein, we first prepared specific antibody. A *trpE-SPT6* gene fusion, encoding SPT6 residues 474 to 860, was constructed by cloning a restriction fragment of *SPT6* into the pATH11 vector (gift of T. J. Koerner, J. Hill, and A. Tzagoloff) to create plasmid pMS1. The TrpE-SPT6 hybrid protein was expressed in *Escherichia coli* HB101 (3), purified by the method of Kleid et al. (19), and used to immunize rabbits. The SPT6-specific antibody was purified from the serum of one rabbit by chromatography on Protein A-Sepharose CL-4B (Pharmacia), followed by affinity chromatography with TrpE-SPT6 protein coupled to CNBr-activated Sepharose (Pharmacia) as the adsorbent according to instructions supplied by the manufacturer. The TrpE-SPT6 hybrid protein used to prepare the column was purified as described by Bhown and Bennett (1).

The SPT6 protein was identified by immunoblot analysis of total yeast proteins. The protein detected was approximately 170 kDa, the size predicted by the *SPT6* DNA sequence. Moreover, the 170-kDa protein was more abun-

* Corresponding author.

```
-396                                                    AGCTGACTGATTCTTTATAGCAATAGGGAGGAAGGCAGTACACATATCAGCAGTGCATGGGCACAACCCGCATTACA

-321  ACGAAACATGATATATCTATATAAAAGAAGATACGGACAGGCGGTCGAGGCTATTACTGTGCCACTATGCGAGTCATGAATGATTTTTAAATTAGTTAGTACCCCTT

-214  TATATAACGTCTTCAATAAAAGTGCGGGAAAATTCGGGTAATATTTTCTATAACTGTCATTTTTTTTTCGTTATGTATGGCAAACCAAATTGACAAGAAACTGGACA

-107  ACGGTTAAAGCAAAAGGAGGAAGAAAGATATTGATTTGTGCTTTGAAGAATAACCCAAGCGAACAGGTCATTAATTGCCTAGAATCTTTACATAGAATTTGCCTTTT

   1  ATG GAA GAG ACG GGA GAT TCG AAG CTG GTC CCT AGG GAC GAG GAA GAA ATA GTA AAT GAC AAC GAT GAA ACT AAA GCG CCT
      Met Glu Glu Thr Gly Asp Ser Lys Leu Val Pro Arg Asp Glu Glu Glu Ile Val Asn Asp Asn Asp Glu Thr Lys Ala Pro    27

  82  AGT GAG GAA GAA GAA GGA GAA GAT GTC TTT GAC TCC TCT GAG GAA GAC GAA GAT ATT GAT GAA GAC GAA GAT GAG GCA AGA
      Ser Glu Glu Glu Glu Gly Glu Asp Val Phe Asp Ser Ser Glu Glu Asp Glu Asp Ile Asp Glu Asp Glu Asp Glu Ala Arg    54

 163  AAA GTG CAA GAA GGC TTC ATT GTC AAC GAT GAT GAT GAA AAT GAA GAC CCA GGA ACA AGT ATT TCC AAA AAA AGA AGA AAA
      Lys Val Gln Glu Gly Phe Ile Val Asn Asp Asp Asp Glu Asn Glu Asp Pro Gly Thr Ser Ile Ser Lys Lys Arg Arg Lys    81

 244  CAT AAA AGA AGA GAA AGA GAA GAA GAT GAT CGA CTA TCC GAA GAT GAT TTG GAT TTG TTA ATG GAG AAT GCT GGT GTT GAA
      His Lys Arg Arg Glu Arg Glu Glu Asp Asp Arg Leu Ser Glu Asp Asp Leu Asp Leu Leu Met Glu Asn Ala Gly Val Glu   108

 325  CGT ACC AAA GCG AGC TCC TCT TCA GGA AAA TTT AAA AGA TTA AAG AGA GTA GGA GAT GAA GGA AAC GCT GCA GAA TCA GAA
      Arg Thr Lys Ala Ser Ser Ser Ser Gly Lys Phe Lys Arg Leu Lys Arg Val Gly Asp Glu Gly Asn Ala Ala Glu Ser Glu   135

 406  AGT GAC AAC GTC GCT GCA TCA AGG CAA GAT TCT ACG TCT AAA TTG GAA GAT TTC TTT TCT GAA GAT GAA GAA GAA GAA GAA
      Ser Asp Asn Val Ala Ala Ser Arg Gln Asp Ser Thr Ser Lys Leu Glu Asp Phe Phe Ser Glu Asp Glu Glu Glu Glu Glu   162

 487  TCT GGT TTA CGC AAC GGT AGG AAC AAT GAA TAT GGG CGT GAT GAA GAA GAC CAT GAA AAT AGG AAC AGA ACA GCT GAC AAA
      Ser Gly Leu Arg Asn Gly Arg Asn Asn Glu Tyr Gly Arg Asp Glu Glu Asp His Glu Asn Arg Asn Arg Thr Ala Asp Lys   189

 568  GGT GGA ATC TTG GAC GAA TTG GAT GAT TTC ATT GAA GAT GAT GAA TTT TCT GAT GAG GAC GAC GAA ACC AGA CAA AGG AGG
      Gly Gly Ile Leu Asp Glu Leu Asp Asp Phe Ile Glu Asp Asp Glu Phe Ser Asp Glu Asp Asp Glu Thr Arg Gln Arg Arg   216

 649  ATC CAA GAA AAG AAG CTT TTA AGA GAA CAA TCC ATC AAA CAA CCT ACA CAG ATT ACT GGT CTA TCG TCG GAT AAG ATT GAC
      Ile Gln Glu Lys Lys Leu Leu Arg Glu Gln Ser Ile Lys Gln Pro Thr Gln Ile Thr Gly Leu Ser Ser Asp Lys Ile Asp   243

 730  GAG ATG TAT GAC ATT TTT GGT GAT GGT CAT GAC TAC GAT TGG GCT TTA GAA ATT GAA AAT GAA GAA CTA GAA AAT GGT AAC
      Glu Met Tyr Asp Ile Phe Gly Asp Gly His Asp Tyr Asp Trp Ala Leu Glu Ile Glu Asn Glu Glu Leu Glu Asn Gly Asn   270

 811  GAC AAC AAT GAA GCT GAA GAA GAA GAG ATC GAT GAA GAA ACT GGT GCT ATA AAA AGC ACC AAG AAA AAG ATA TCT CTA CAA
      Asp Asn Asn Glu Ala Glu Glu Glu Glu Ile Asp Glu Glu Thr Gly Ala Ile Lys Ser Thr Lys Lys Lys Ile Ser Leu Gln   297

 892  GAC ATT TAT GAT TTA GAG GAT TTG AAA AAA AAC TTG ATG ACT GAA GGA GAC ATG AAA ATT AGA AAG ACA GAT ATT CCA GAA
      Asp Ile Tyr Asp Leu Glu Asp Leu Lys Lys Asn Leu Met Thr Glu Gly Asp Met Lys Ile Arg Lys Thr Asp Ile Pro Glu   324

 973  AGA TAT CAA GAA TTA AGA GCA GGT ATT ACT GAC TAC GGA AAT ATG TCA TCG GAG GAT CAA GAA TTA GAA AGA AAC TGG ATA
      Arg Tyr Gln Glu Leu Arg Ala Gly Ile Thr Asp Tyr Gly Asn Met Ser Ser Glu Asp Gln Glu Leu Glu Arg Asn Trp Ile   351

1054  GCA GAA AAA ATT TCT GTG GAT AAG AAC TTC GAT GCC AAT TAT GAT CTC ACT GAA TTT AAA GAA GCA ATT GGG AAC GCA ATC
      Ala Glu Lys Ile Ser Val Asp Lys Asn Phe Asp Ala Asn Tyr Asp Leu Thr Glu Phe Lys Glu Ala Ile Gly Asn Ala Ile   378

1135  AAA TTT ATC ACC AAA GAA AAC TTG GAA GTC CCT TTT ATA TAT GCT TAC CGT CGT AAC TAT ATT TCC TCA AGA GAA AAA GAT
      Lys Phe Ile Thr Lys Glu Asn Leu Glu Val Pro Phe Ile Tyr Ala Tyr Arg Arg Asn Tyr Ile Ser Ser Arg Glu Lys Asp   405

1216  GGG TTT CTT TTG ACT GAA GAT GAC CTT TGG GAT ATA GTT AGC CTT GAC ATC GAA TTT CAC AGT CTT GTG AAC AAA AAG GAT
      Gly Phe Leu Leu Thr Glu Asp Asp Leu Trp Asp Ile Val Ser Leu Asp Ile Glu Phe His Ser Leu Val Asn Lys Lys Asp   432

1297  TAT GTC CAG AGA TTT TAT GCA GAA TTA CAT ATC GAT GAT CCT ATT GTC ACT GAA TAC TTC AAA AAT CAG AAT ACT GCA TCT
      Tyr Val Gln Arg Phe Tyr Ala Glu Leu His Ile Asp Asp Pro Ile Val Thr Glu Tyr Phe Lys Asn Gln Asn Thr Ala Ser   459
                                                          EcoRI
1378  ATT GCG GAA CTG AAT TCT CTA CAG GAT ATT TAT GAC TAC CTA GAA TTC AAA TAT GCC AAC GAA ATC AAT GAA ATG TTT ATA
      Ile Ala Glu Leu Asn Ser Leu Gln Asp Ile Tyr Asp Tyr Leu Glu Phe Lys Tyr Ala Asn Glu Ile Asn Glu Met Phe Ile   486

1459  AAC CAC ACT GGA AAG ACT GGT AAG AAA CAT TTG AAA AAT TCC AGT TAT GAA AAA TTT AAA GCT AGT CCT CTT TAT CAA GCG
      Asn His Thr Gly Lys Thr Gly Lys Lys His Leu Lys Asn Ser Ser Tyr Glu Lys Phe Lys Ala Ser Pro Leu Tyr Gln Ala   513

1540  GTT AGT GAT ATT GGT ATA TCA GCT GAG GAT GTT GGT GAA AAT ATC AGT TCC CAG CAT CAA ATC CAC CCT CCC GTA GAT CAT
      Val Ser Asp Ile Gly Ile Ser Ala Glu Asp Val Gly Glu Asn Ile Ser Ser Gln His Gln Ile His Pro Pro Val Asp His   540

1621  CCA AGT TCC AAA CCA GTA GAA GTG ATA GAA TCT ATA TTG AAT GCA AAC AGC GGT GAT TTG CAA GTC TTT ACG TCC AAT ACT
      Pro Ser Ser Lys Pro Val Glu Val Ile Glu Ser Ile Leu Asn Ala Asn Ser Gly Asp Leu Gln Val Phe Thr Ser Asn Thr   567

1702  AAG CTG GCA ATT GAT ACG GTC CAA AAA TAC TAC TCT TTG GAA TTG TCT AAA AAT ACA AAA ATA AGG GAA AAA GTT AGA TCC
      Lys Leu Ala Ile Asp Thr Val Gln Lys Tyr Tyr Ser Leu Glu Leu Ser Lys Asn Thr Lys Ile Arg Glu Lys Val Arg Ser   594

1783  GAT TTT TCC AAA TAT TAT CTG GCT GAC GTT GTG TTA ACT GCT AAA GGT AAA AAA GAA ATT CAA AAG GGA TCT CTG TAT GAG
      Asp Phe Ser Lys Tyr Tyr Leu Ala Asp Val Val Leu Thr Ala Lys Gly Lys Lys Glu Ile Gln Lys Gly Ser Leu Tyr Glu   621

1864  GAC ATA AAA TAT GCC ATC AAT AGA ACT CCA ATG CAC TTC CGT AGG GAT CCA GAC GTT TTT TTG AAA ATG GTC GAG GCT GAG
      Asp Ile Lys Tyr Ala Ile Asn Arg Thr Pro Met His Phe Arg Arg Asp Pro Asp Val Phe Leu Lys Met Val Glu Ala Glu   648

1945  TCT TTG AAC CTG CTC AGT GTT AAG TTA CAC ATG TCG TCA CAA GCC CAA TAT ATA GAG CAT TTA TTC CAA ATT GCA CTT GAA
      Ser Leu Asn Leu Leu Ser Val Lys Leu His Met Ser Ser Gln Ala Gln Tyr Ile Glu His Leu Phe Gln Ile Ala Leu Glu   675

2026  ACT ACC AAT ACC TCG GAC ATC GCA ATA GAA TGG AAT AAC TTC CGT AAA CTG GCA TTC AAC CAA GCG ATG GAC AAG ATT TTC
      Thr Thr Asn Thr Ser Asp Ile Ala Ile Glu Trp Asn Asn Phe Arg Lys Leu Ala Phe Asn Gln Ala Met Asp Lys Ile Phe   702

2107  CAG GAT ATA TCT CAA GAA GTC AAA GAC AAT TTA ACA AAA AAT TGT CAA AAA TTG GTA GCC AAG ACT GTT CGC CAT AAG TTT
      Gln Asp Ile Ser Gln Glu Val Lys Asp Asn Leu Thr Lys Asn Cys Gln Lys Leu Val Ala Lys Thr Val Arg His Lys Phe   729
```

```
2188  ATG ACA AAA TTA GAC CAG GCT CCA TTC ATT CCT AAT GTC AGG GAT CCA AAA ATT CCA AAA ATC TTA TCT TTA ACC TGT GGA
      Met Thr Lys Leu Asp Gln Ala Pro Phe Ile Pro Asn Val Arg Asp Pro Lys Ile Pro Lys Ile Leu Ser Leu Thr Cys Gly    756

2269  CAG GGT AGA TTC GGA GCC GAC GCT ATA ATT GCT GTC TAC GTC AAC AGA AAG GGT GAT TTT ATA AGA GAT TAC AAG ATT GTC
      Gln Gly Arg Phe Gly Ala Asp Ala Ile Ile Ala Val Tyr Val Asn Arg Lys Gly Asp Phe Ile Arg Asp Tyr Lys Ile Val    783

2350  GAC AAT CCA TTT GAT AAG ACG AAT CCT GAA AAA TTT GAA GAC ACC TTG GAT AAT ATC ATT CAA AGC TGT CAA CCG AAT GCC
      Asp Asn Pro Phe Asp Lys Thr Asn Pro Glu Lys Phe Glu Asp Thr Leu Asp Asn Ile Ile Gln Ser Cys Gln Pro Asn Ala    810

2431  ATC GGA ATC AAT GGC CCT AAC CCA AAG ACT CAA AAA TTT TAC AAA AGA TTA CAA GAA GTT CTA CAT AAG AAG CAA ATC GTC
      Ile Gly Ile Asn Gly Pro Asn Pro Lys Thr Gln Lys Phe Tyr Lys Arg Leu Gln Glu Val Leu His Lys Lys Gln Ile Val    837
                                                                                  EcoRI
2512  GAC AGT AGA GGA CAT ACT ATT CCA ATC ATT TAC GTT GAG GAC GAA GTC GCT ATC CGT TAT CAG AAT TCC GAA AGA GCT GCT
      Asp Ser Arg Gly His Thr Ile Pro Ile Ile Tyr Val Glu Asp Glu Val Ala Ile Arg Tyr Gln Asn Ser Glu Arg Ala Ala    864

2593  CAA GAA TTC CCT AAT AAA CCT CCT CTA GTT AAA TAC TGT ATC GCC TTG GCG CGC TAT ATG CAT TCC CCA TTG TTG GAA TAT
      Gln Glu Phe Pro Asn Lys Pro Pro Leu Val Lys Tyr Cys Ile Ala Leu Ala Arg Tyr Met His Ser Pro Leu Leu Glu Tyr    891

2674  GCT AAT TTA ACA AGT GAA GAA GTG AGA TCA TTG TCA ATT CAT CCA CAC CAA AAT CTG TTA TCC TCA GAA CAA TTG AGT TGG
      Ala Asn Leu Thr Ser Glu Glu Val Arg Ser Leu Ser Ile His Pro His Gln Asn Leu Leu Ser Ser Glu Gln Leu Ser Trp    918

2755  GCT CTT GAA ACC GCT TTC GTT GAT ATT GTC AAC CTG GTA AGT GTT GAA GTT AAC AAA GCC ACA GAT AAT AAT TAC TAC GCT
      Ala Leu Glu Thr Ala Phe Val Asp Ile Val Asn Leu Val Ser Val Glu Val Asn Lys Ala Thr Asp Asn Asn Tyr Tyr Ala    945

2836  AGT GCG CTG AAA TAC ATC TCT GGC TTT GGA AAA CGT AAA GCT ATT GAT TTC TTA CAG TCC CTT CAA AGG CTA AAT GAA CCA
      Ser Ala Leu Lys Tyr Ile Ser Gly Phe Gly Lys Arg Lys Ala Ile Asp Phe Leu Gln Ser Leu Gln Arg Leu Asn Glu Pro    972

2917  TTA CTG GCT CGT CAA CAA TTA ATT ACT CAT AAC ATT CTT CAC AAG ACT ATT TTT ATG AAT TCC GCG GGA TTC CTC TAT ATC
      Leu Leu Ala Arg Gln Gln Leu Ile Thr His Asn Ile Leu His Lys Thr Ile Phe Met Asn Ser Ala Gly Phe Leu Tyr Ile    999

2998  TCA TGG AAT GAA AAA AGA CAA AAA TAC GAA GAT TTG GAA CAT GAT CAA CTA GAT AGC ACT AGA ATT CAT CCA GAA GAC TAC
      Ser Trp Asn Glu Lys Arg Gln Lys Tyr Glu Asp Leu Glu His Asp Gln Leu Asp Ser Thr Arg Ile His Pro Glu Asp Tyr   1026

3079  CAT TTG GCC ACC AAG GTT GCC GCT GAT GCT TTA GAA TAC GAT CCT GAT ACT ATT GCC GAA AAA GAA GAA CAG GGG ACT ATG
      His Leu Ala Thr Lys Val Ala Ala Asp Ala Leu Glu Tyr Asp Pro Asp Thr Ile Ala Glu Lys Glu Glu Gln Gly Thr Met   1053

3160  AGT GAA TTC ATT GAA CTG TTG AGA GAA GAT CCT GAC CGT AGA GCT AAA CTA GAA TCA CTA AAT CTA GAA TCA TAC GCA GAA
      Ser Glu Phe Ile Glu Leu Leu Arg Glu Asp Pro Asp Arg Arg Ala Lys Leu Glu Ser Leu Asn Leu Glu Ser Tyr Ala Glu   1080

3241  GAA CTT GAG AAG AAT ACC GGA TTA AGA AAA CTT AAT AAT CTA AAT ACA ATT GTC CTT GAA TTG TTG GAT GGA TTT GAA GAA
      Glu Leu Glu Lys Asn Thr Gly Leu Arg Lys Leu Asn Asn Leu Asn Thr Ile Val Leu Glu Leu Leu Asp Gly Phe Glu Glu   1107

3322  TTG AGA AAT GAC TTT CAT CCT TTG CAA GGT GAT GAA ATT TTC CAA AGT TTG ACT GGT GAG TCT GAA AAG ACG TTT TTC AAG
      Leu Arg Asn Asp Phe His Pro Leu Gln Gly Asp Glu Ile Phe Gln Ser Leu Thr Gly Glu Ser Glu Lys Thr Phe Phe Lys   1134

3403  GGT AGT ATT ATT CCA GTC AGA GTA GAA AGA TTC TGG CAC AAC GAT ATA ATT TGC ACT ACA AAC TCT GAA GTT GAA TGT GTA
      Gly Ser Ile Ile Pro Val Arg Val Glu Arg Phe Trp His Asn Asp Ile Ile Cys Thr Thr Asn Ser Glu Val Glu Cys Val   1161

3484  GTA AAT GCT CAA CGT CAC GCA GGT GCA CAA TTA AGA AGA CCT GCA AAT GAA ATA TAC GAA ATT GGT AAA ACA TAT CCA GCA
      Val Asn Ala Gln Arg His Ala Gly Ala Gln Leu Arg Arg Pro Ala Asn Glu Ile Tyr Glu Ile Gly Lys Thr Tyr Pro Ala   1188

3565  AAG GTG ATA TAT ATT GAC TAT GCT AAT ATT ACT GCA GAA GTT TCC TTA TTA GAT CAT GAT GTC AAA CAG CAA TAT GTT CCA
      Lys Val Ile Tyr Ile Asp Tyr Ala Asn Ile Thr Ala Glu Val Ser Leu Leu Asp His Asp Val Lys Gln Gln Tyr Val Pro   1215

3646  ATA AGC TAC AGT AAA GAT CCT TCC ATT TGG GAC TTG AAA CAA GAA CTG GAA GAT GCC GAA GAG GAG AGG AAA TTG ATG ATG
      Ile Ser Tyr Ser Lys Asp Pro Ser Ile Trp Asp Leu Lys Gln Glu Leu Glu Asp Ala Glu Glu Glu Arg Lys Leu Met Met   1242

3727  GCA GAA GCC CGT GCA AAG AGA ACA CAT CGT GTT ATC AAT CAT CCT TAC TAT TTC CCT TTC AAC GGC AGA CAG GCT GAG GAT
      Ala Glu Ala Arg Ala Lys Arg Thr His Arg Val Ile Asn His Pro Tyr Tyr Phe Pro Phe Asn Gly Arg Gln Ala Glu Asp   1269

3808  TAC TTA AGG AGT AAA GAA CGT GGT GAA TTC GTG ATC AGA CAG TCT AGC CGA GGT GAT GAC CAC TTG GTT ATC ACC TGG AAA
      Tyr Leu Arg Ser Lys Glu Arg Gly Glu Phe Val Ile Arg Gln Ser Ser Arg Gly Asp Asp His Leu Val Ile Thr Trp Lys   1296

3889  TTG GAT AAG GAT TTG TTT CAA CAT ATT GAT ATC CAA GAA TTA GAA AAA GAA AAT CCT TTG GCT TTA GGT AAA GTC TTG ATT
      Leu Asp Lys Asp Leu Phe Gln His Ile Asp Ile Gln Glu Leu Glu Lys Glu Asn Pro Leu Ala Leu Gly Lys Val Leu Ile   1323

3970  GTC GAC AAT CAG AAA TAC AAT GAT TTA GAC CAG ATC ATT GTA GAA TAT CTT CAA AAC AAG GTA AGG CTC TTG AAT GAA ATG
      Val Asp Asn Gln Lys Tyr Asn Asp Leu Asp Gln Ile Ile Val Glu Tyr Leu Gln Asn Lys Val Arg Leu Leu Asn Glu Met   1350

4051  ACA TCT AGT GAA AAA TTC AAA AGC GGT ACT AAG AAA GAT GTG GTC AAG TTT ATT GAA GAC TAC TCT AGA GTG AAT CCA AAT
      Thr Ser Ser Glu Lys Phe Lys Ser Gly Thr Lys Lys Asp Val Val Lys Phe Ile Glu Asp Tyr Ser Arg Val Asn Pro Asn   1377

4132  AAG TCT GTG TAC TAT TTC AGT TTG AAC CAC GAT AAC CCT GGT TGG TTT TAC TTG ATG TTC AAG ATT AAC GCA AAT AGC AAA
      Lys Ser Val Tyr Tyr Phe Ser Leu Asn His Asp Asn Pro Gly Trp Phe Tyr Leu Met Phe Lys Ile Asn Ala Asn Ser Lys   1404

4213  TTA TAC ACA TGG AAT GTG AAA TTA ACG AAC ACT GGT TAT TTC CTG GTA AAC TAC AAT TAT CCA AGT GTT ATC CAG CTT TGT
      Leu Tyr Thr Trp Asn Val Lys Leu Thr Asn Thr Gly Tyr Phe Leu Val Asn Tyr Asn Tyr Pro Ser Val Ile Gln Leu Cys   1431

4294  AAT GGT TTT AAG ACG CTT CTA AAA TCT AAC AGT AGT AAG AAT AGA ATG AAC AAC TAC CGT TAG ATGCGTATGTAGTGTCCATTGTT
      Asn Gly Phe Lys Thr Leu Leu Lys Ser Asn Ser Ser Lys Asn Arg Met Asn Asn Tyr Arg ***                        1451

4380  ATTATTAATTTTATTATTACTTTGACCATTGTTTAGGTAAGGTATGTAAACTAGAACGCTTCAACTGTATTTTTTACGTAATTATTTGCTGCATGTTTAGTTTCGCA

4487  AAGTTTGTACCGTGACGAACATAGCACGAACACAACATTTAATAGAAATAGCTTGGGAACCAAAATTTTAGAAATTTTTTTCACTAATGAGAAGGAAAGCGTTGCAA

4594  AGTGGCGACTATTTTGGTTCTTCAAATATTTGCAGGAAATAAATAATTTACTGTGCATTTCTGATTACATTTAGGATGTAAATTGACTTAGTAGACGTTCCA
```

FIG. 1. Nucleotide sequence of the *SPT6* gene (GenBank accession number M34391) and predicted amino acid sequence of its gene product. Nucleotides are numbered on the left. Amino acids are numbered on the right. Asterisks indicate the termination codon. Acidic residues are underlined. The *Eco*RI restriction sites used to construct the *trpE-SPT6* gene fusion used for antibody preparation are indicated. Additional *Eco*RI sites are located at positions 1389, 2596, 2973, 3059, 3163, and 3832. The *Hin*dIII site used to construct a null mutation (6) is at position 661. The complete sequence was determined on both strands with the M13mp18 and M13mp19 vectors (27). In four cases the restriction fragments were digested with Bal31 nuclease (IBI) prior to cloning, and in six cases 20-base oligomers (Biopolymers Laboratory, Department of Genetics, Harvard Medical School) complementary to the adjacent *SPT6* sequence were used as primers.

dant in yeast cells carrying the *SPT6* gene on a multicopy plasmid (Fig. 2, compare lanes A and B).

Nuclear localization of the SPT6 protein. The distribution of the SPT6 protein in yeast cells was determined by indirect immunofluorescence microscopy. Initially, we attempted to localize the SPT6 protein by using the affinity-purified SPT6-specific antibody. This antibody stained cells faintly throughout the cell (data not shown). However, because an *spt6* null mutation is lethal (6, 25), a yeast strain lacking SPT6 protein was not available as a control to determine whether the staining was due to nonspecific binding of the antibody.

To circumvent this problem, we used an epitope addition method (12, 23). A hybrid gene encoding an antigenic epitope from the hemagglutinin of influenza virus (HA1 [12, 27]) fused to the SPT6 protein was constructed by using the Muta-Gene M13 in vitro mutagenesis kit (Bio-Rad Laboratories). The template for the mutagenesis was pMS26, which contains the 1.8-kb *Hin*dIII fragment from pCC11 (6) cloned into the *Hin*dIII site of M13mp19 (28). The primer used for the mutagenesis was a 69-base oligonucleotide (Biopolymers Laboratory, Department of Genetics, Harvard Medical School) including the 17 nucleotides 5′ to the *SPT6* initiation codon, the initiation codon, the 27 nucleotides encoding the HA1 epitope, and then the 22 nucleotides 3′ to the *SPT6* initiation codon. The mutated 1.8-kb *Hin*dIII fragment was used to create pMS33, which encodes the HA1 epitope fused to the complete SPT6 protein on a centromere-containing plasmid. The hybrid protein was functional, as judged by its ability to complement the temperature-sensitive lethality and Spt$^-$ phenotype caused by an *spt6-140* mutation (32) (data not shown).

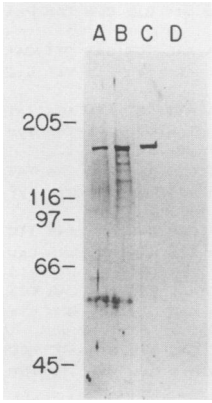

FIG. 2. Immunoblot analysis of the SPT6 protein. Total protein (40 μg in lanes A and B and 25 μg in lanes C and D) was prepared as described previously (4), separated by electrophoresis in 7.5% polyacrylamide, and electroblotted to nitrocellulose. Protein samples were prepared from (lane A) wild-type yeast strain MCRY1093, (lane B) an *spt6* mutant strain carrying *SPT6* on a multicopy plasmid (MCRY237), (lane C) wild-type yeast cells carrying the *HA1-SPT6* hybrid gene on a centromere-containing plasmid (MS147), and (lane D) wild-type yeast cells carrying the *SPT6* gene on a centromere-containing plasmid (MS146). SPT6-specific antibody affinity-purified from rabbit and diluted 1:20 was used with the Promega Protoblot system at a 1:7,500 dilution of rabbit immunoglobulin G (IgG)-specific antibody affinity-purified from goat serum and conjugated to alkaline phosphatase (lanes A and B). HA1-specific antibody was diluted 1:100 and used with a 1:7,500 dilution of mouse IgG-specific antibody affinity-purified from goat serum and conjugated to alkaline phosphatase (lanes C and D). The minor bands in lane B were not found consistently and are presumably degradation products. Numbers at right mark positions of protein standards (in kilodaltons).

To study the HA1-SPT6 hybrid protein, we used a specific monoclonal antibody against the HA1 epitope, called 12CA5, that was prepared and characterized previously (27). The specificity of the anti-HA1 antibody for the HA1-SPT6 hybrid protein in yeast cells was determined first by immunoblot analysis. The epitope-tagged SPT6 protein detected by the HA1-specific antibody (Fig. 2, lane C) comigrated with the wild-type 170-kDa SPT6 protein detected by SPT6-specific antibodies (Fig. 2, lanes A and B). No staining by anti-HA1 antibody was observed in a strain lacking the HA1 epitope (Fig. 2, lane D).

To determine the cellular location of the SPT6 protein, tetraploid cells (Table 1), whose large size facilitated microscopic analysis, were prepared for immunofluorescence microscopy. Experiments with the HA1-specific antibody and a strain containing the HA1-SPT6 fusion protein demonstrated that the epitope-tagged SPT6 protein was located in the nucleus of cells, as evident from the coincident fluorescein and DAPI (4′,6′-diamidino-2-phenylindole dihydrochloride) staining (Fig. 3A to F). In the control experiment, no staining was detected in cells lacking the HA1 epitope (Fig. 3G to I). Since the HA1-SPT6 fusion protein provided SPT6 function, it is very likely that the wild-type SPT6 protein is also located in the nucleus.

To address the possibility that the HA1 epitope itself directs proteins to the nucleus, we determined the cellular location of an HA1–β-galactosidase fusion protein. Previous

TABLE 1. Yeast strains

Strain	Genotype[a]
MCRY1093	*MATa his4-539 lys2-801 ura3-52*
MCRY237	*MATa his4-539 ura3-52 ssn20-1* [pCE204]
MS146	*MATa his4-912δ lys2-128δ ura3-52 trp1Δ63* [pCC11]
MS147	*MATa his4-912δ lys2-128δ ura3-52 trp1Δ63* [pMS33]
MS149	*MATa/MATa ura3-52/ura3-52 his4-912δ/his4-912δ lys2-128δ/lys2-128δ leu2Δ1/+ trp1Δ63/+* [pMS33]
MS150	*MATα/MATα ura3-52/ura3-52 his4-912δ/his4-912δ lys2-128δ/lys2-128δ leu2Δ1/+ trp1Δ63/+spt6-140/+*
MS148	MS149 × MS150
BM330	*MATα/MATα ura3-52/ura3-52 his4-912δ/his4-912δ lys2-128δ/lys2-128δ trp1Δ63/trp1Δ63*
BM331	*MATa/MATa ura3-52/ura3-52 his4-912δ/his4-912δ lys2-128δ/lys2-128δ trp1Δ63/trp1Δ63*
BM339	BM330 × BM331
MS157	BM339 [pMS35]

[a] Plasmids are indicated in brackets. Diploid strains homozygous at the *MAT* locus were generated by UV-irradiating diploid cells (300 ergs/mm^2) and then screening for diploids that could mate.

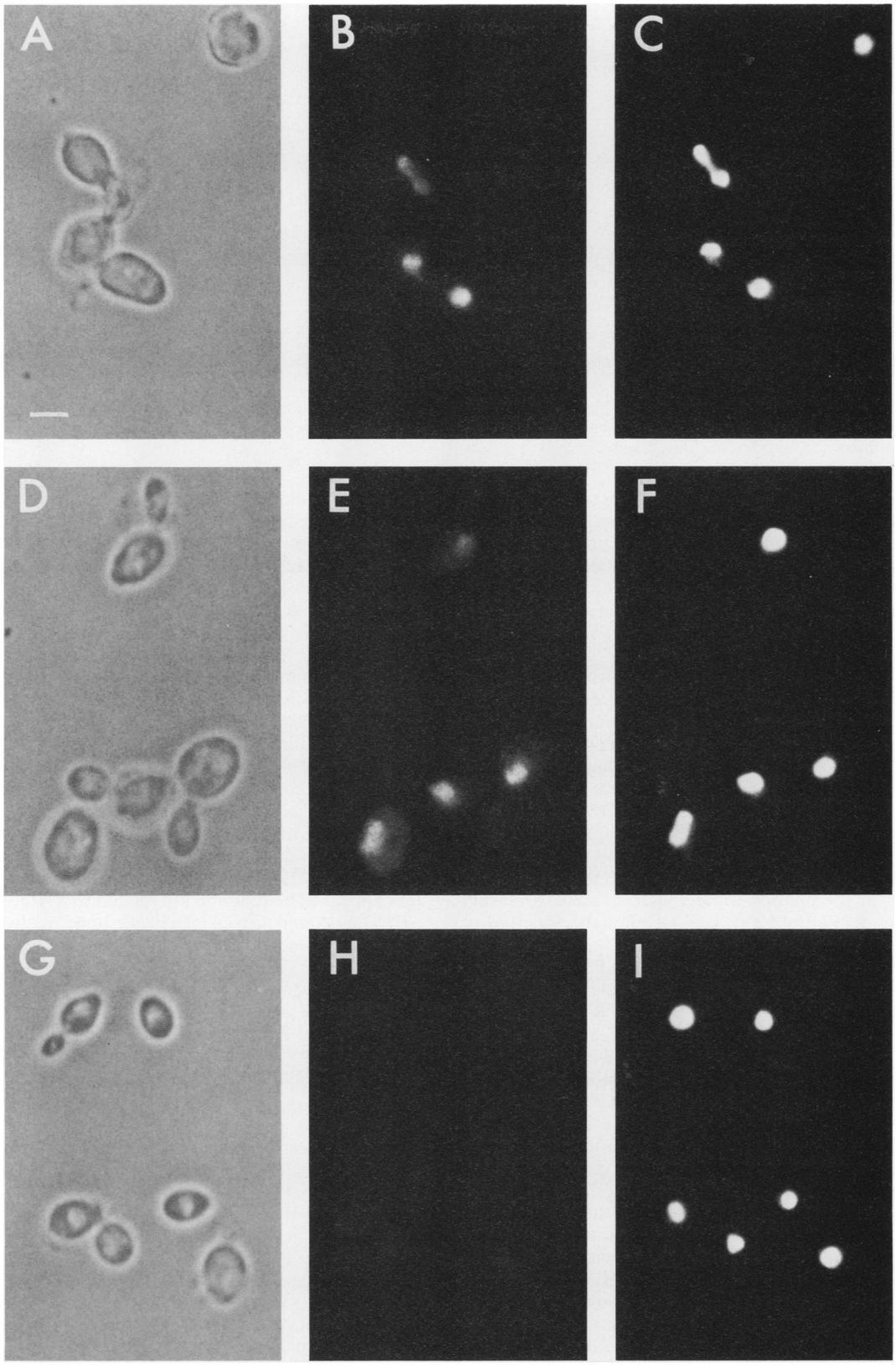

FIG. 3. Nuclear localization of SPT6 protein by indirect immunofluorescence. A tetraploid yeast strain containing the HA1-SPT6 fusion protein (MS148) (A to F) and a congenic control strain lacking the HA1 epitope (BM339) (G to I) were prepared essentially as described by Kilmartin and Adams (18), except that cell cultures were fixed in 4% paraformaldehyde (Fisher). The cells were stained with HA1-specific antibody diluted 1:400 and affinity-purified mouse IgG-specific $F(ab')_2$ fragment conjugated to fluorescein diluted 1:2,000. Micrographs shown are phase contrast (A, D, and G), fluorescein fluorescence (B, E, and H), and DAPI fluorescence, which indicates DNA (C, F, and I). Cells were examined with a Zeiss Photomicroscope III equipped for epi-illumination fluorescence with a Zeiss Neofluar 63× lens, N.A. 1.25. Cells were photographed with Kodak TMAX-400 film. Exposure times of 40 s were used for the immunofluorescence micrographs of the experimental and control samples. Bar, 5 μm.

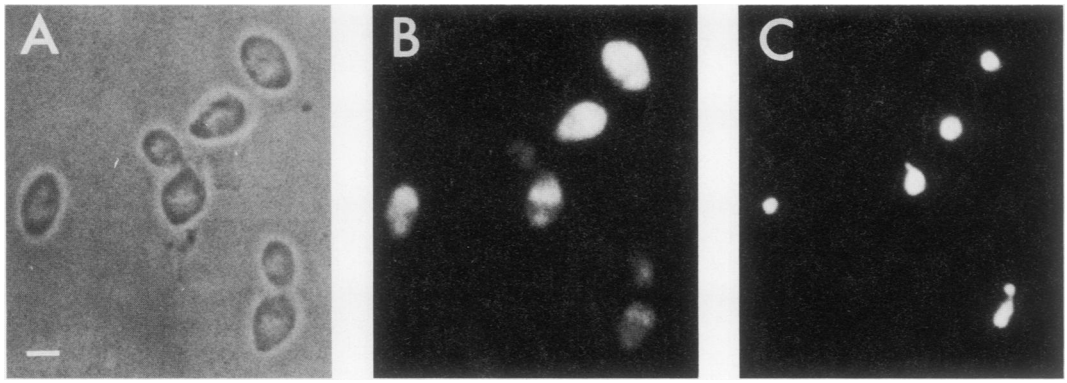

FIG. 4. Localization of HA1–β-galactosidase by indirect immunofluorescence. Tetraploid yeast cells containing the HA1–β-galactosidase fusion protein (MS157) were stained with HA1-specific antibody as described in the legend to Fig. 3. An exposure time of 25 s was used for the immunofluorescence micrograph. Bar, 5 μm.

work has shown that β-galactosidase expressed in yeast cells is distributed throughout the cytoplasm and nucleus (13). We first constructed an *HA1-lacZ* hybrid gene under control of the *ADH1* promoter. This construct was made by cloning the 1.5-kb *Bam*HI-*Sal*I fragment from pAD5, which contains the *ADH1* promoter and the HA1 epitope-coding sequences (gift of J. Field, L. Rogers, and M. Wigler), into YEp353, which contains the coding region of *E. coli lacZ* (24), to create pMS35, which was then used to transform yeast cells. Immunofluorescence experiments demonstrated that the HA1–β-galactosidase fusion protein was located throughout yeast cells (Fig. 4). In the control experiment, no staining was detected in a strain that did not contain the epitope-tagged β-galactosidase fusion protein (data not shown). Our observation that the HA1 epitope did not direct β-galactosidase to the nucleus is consistent with the fact that the HA1 amino acid sequence (YPYDVPDYA) does not resemble known nuclear signal sequences (9) and with the observation that addition of the epitope to adenylate cyclase, a plasma membrane-associated protein, did not alter the function of this enzyme in vivo (12).

The amino terminus of the predicted SPT6 protein included three regions that may serve as nuclear localization signals (see reference 9 for review): KLVPR (residues 8 to 12), KKRRKHKRR (residues 77 to 85), and KRLKRV (residues 120 to 125). The actual role of the these regions will have to be determined experimentally, since a variety of primary sequences have been shown to function as nuclear localization signals (9).

Nuclear localization of the SPT6 protein is consistent with the *spt6* mutant phenotypes characterized previously. Genetic analysis has demonstrated that *SPT6* affects transcription of a variety of genes and suggested that the SPT6 protein may act negatively at a downstream promoter element (8, 25). SPT6 protein may interact directly with DNA via a previously unidentified DNA-binding motif. Alternatively, SPT6 protein may play a role in chromatin assembly or modification. Many studies have correlated alterations in chromatin structure with transcriptional activation (for reviews, see references 22 and 29). Like SPT6, proteins believed to interact with chromatin have extremely acidic regions (10). Furthermore, the *SPT6* gene dosage effects are similar to those caused by histone genes. Either increased or decreased dosage of some histone genes also suppresses δ insertion mutations (5) and deletions of the *SUC2* upstream regulatory region (J. Hirschhorn and F. Winston, unpublished observations). Han and Grunstein (14) have shown that, like *spt6* mutations (25), nucleosome depletion restores mRNA expression from alleles with upstream activation sequence deletions.

One approach we are taking to determine the function of *SPT6* is to identify genes encoding proteins that are functionally related to SPT6. Since *SPT6* is an essential gene, it is conceivable that mutations in functionally related genes would cause lethality when combined with temperature-sensitive *spt6* mutations. Two candidate genes are *SPT4* and *SPT5*, which were also identified in genetic selections for extragenic suppressors of Ty insertion mutations (11, 32). The combination of any two of the missense mutations *spt4-3*, *spt5-194*, and *spt6-140* causes lethality (32). Moreover, some combinations of *spt4*, *spt5*, and *spt6* alleles exhibit unlinked noncomplementation (M. S. Swanson and F. Winston, unpublished observations), a phenotype observed for mutations in some proteins that are known to interact (15, 31). By genetic and molecular analysis of genes that appear to be functionally related to *SPT6*, such as *SPT4*, *SPT5*, and some histone genes, we hope to elucidate the role of *SPT6* in gene expression in yeast.

We thank John Celenza for advice and assistance in sequencing the *SPT6* gene and preparation of the TrpE-SPT6 antigen and antiserum and Laura Davis and Linda Marshall-Carlson for helpful suggestions regarding immunofluorescence methods.

This work was supported by Public Health Service grants GM34095 to M.C. and GM32967 to F.W. from the National Institutes of Health (NIH). At Columbia University, M.S. was supported by NIH training grant T32-GM07088 to the Department of Genetics and Development. At Harvard Medical School, M.S. was supported by NIH training grants to the Cellular and Developmental Biology Program (T32-GM07226) and the Genetics Program (T32-GM07196) and by the Markey Charitable Trust.

LITERATURE CITED

1. **Bhown, A. S., and J. C. Bennett.** 1983. High-sensitivity sequence analysis of proteins recovered from sodium dodecyl sulfate gels. Methods Enzymol. **91**:450–455.
2. **Blumenthal, D. K., A. Edelman, and E. G. Krebs.** 1987. Protein serine/threonine kinases. Annu. Rev. Biochem. **56**:567–613.
3. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *E. coli*. J. Mol. Biol. **41**:458–472.
4. **Celenza, J. L., and M. Carlson.** 1986. A yeast gene that is essential for release from glucose repression encodes a protein kinase. Science **233**:1175–1180.
5. **Clark-Adams, C. D., D. Norris, M. A. Osley, J. S. Fassler, and F. Winston.** 1988. Changes in histone gene dosage alter transcription in yeast. Genes Dev. **2**:150–159.

6. **Clark-Adams, C. D., and F. Winston.** 1987. The *SPT6* gene is essential for growth and is required for δ-medicated transcription in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7**:679–686.
7. **Denis, C. L.** 1984. Identification of new genes involved in the regulation of yeast alcohol dehydrogenase II. Genetics **108**:845–858.
8. **Denis, C. L., and T. Malvar.** 1990. The *CCR4* gene from *Saccharomyces cerevisiae* is required for both nonfermentative and *spt*-mediated gene expression. Genetics **124**:283–291.
9. **Dingwall, C., and R. A. Laskey.** 1986. Protein import into the cell nucleus. Annu. Rev. Cell Biol. **2**:367–390.
10. **Earnshaw, W. C.** 1987. Anionic regions in nuclear proteins. J. Cell Biol. **105**:1479–1482.
11. **Fassler, J. S., and F. Winston.** 1988. Isolation and analysis of a novel class of suppressor of Ty insertion mutations in *Saccharomyces cerevisiae*. Genetics **118**:203–212.
12. **Field, J., J.-I. Nikawa, D. Broek, B. MacDonald, L. Rodgers, I. A. Wilson, R. A. Lerner, and M. Wigler.** 1988. Purification of a *RAS*-responsive adenylyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method. Mol. Cell. Biol. **8**:2159–2165.
13. **Hall, M. N., L. Hereford, and I. Herskowitz.** 1984. Targeting of *E. coli* β-galactosidase to the nucleus in yeast. Cell **36**:1057–1065.
14. **Han, M., and M. Grunstein.** 1988. Nucleosome loss activates yeast downstream promoters *in vivo*. Cell **55**:1137–1145.
15. **Hayes, T. S., R. Deuring, B. Robertson, M. Prout, and M. T. Fuller.** 1989. Interacting proteins identified by genetic interactions: a missense mutation in α-tubulin fails to complement alleles of the testis-specific β-tubulin gene of *Drosophila melanogaster*. Mol. Cell. Biol. **9**:875–884.
16. **Hope, I. A., S. Mahadevan, and K. Struhl.** 1988. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. Nature (London) **333**:635–640.
17. **Hope, I. A., and K. Struhl.** 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell **46**:885–894.
18. **Kilmartin, J. V., and A. E. M. Adams.** 1984. Structural rearrangements of tubulin and actin during the cell cycle of the yeast *Saccharomyces*. J. Cell Biol. **98**:922–933.
19. **Kleid, D. G., D. Yansura, B. Small, D. Dowbenko, D. M. Moore, M. J. Gibman, P. D. McKercher, D. O. Morgan, B. H. Robertson, and H. L. Bachrach.** 1981. Cloned viral protein vaccine for foot-and-mouth disease: responses in cattle and swine. Science **214**:1125–1129.
20. **Kuenzel, E. A., J. A. Mulligan, J. Sommercorn, and E. G. Krebs.** 1987. Substrate specificity determinants for casein kinase II as deduced from studies with synthetic peptides. J. Biol. Chem. **262**:9136–9140.
21. **Ma, J., and M. Ptashne.** 1987. Deletion analysis of *GAL4* defines two transcriptional activating segments. Cell **48**:847–853.
22. **Mathis, D., P. Oudet, and P. Chambon.** 1980. Structure of transcribing chromatin. Prog. Nucleic Acids Res. Mol. Biol. **24**:1–55.
23. **Munro, S., and H. R. B. Pelham.** 1987. A C-terminal signal prevents secretion of luminal ER proteins. Cell **48**:899–907.
24. **Myers, A. M., A. Tzagoloff, D. M. Kinney, and C. J. Lusty.** 1986. Yeast shuttle and integrative vectors with multiple cloning sites suitable for construction of *lacZ* fusions. Gene **45**:299–310.
25. **Neigeborn, L., J. L. Celenza, and M. Carlson.** 1987. *SSN20* is an essential gene with mutant alleles that suppress defects in *SUC2* transcription in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7**:672–678.
26. **Neigeborn, L., K. Rubin, and M. Carlson.** 1986. Suppressors of *snf2* mutations restore invertase derepression and cause temperature-sensitive lethality in yeast. Genetics **112**:741–753.
27. **Niman, H. L., R. A. Houghten, L. E. Walker, R. A. Reisfeld, I. A. Wilson, J. M. Hogle, and R. A. Lerner.** 1983. Generation of protein-reactive antibodies by short peptides is an event of high frequency: implications for the structural basis of immune recognition. Proc. Natl. Acad. Sci. USA **80**:4949–4953.
28. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.
29. **Pederson, D. S., F. Thoma, and R. T. Simpson.** 1986. Core particle, fiber, and transcriptionally active chromatin structure. Annu. Rev. Cell Biol. **2**:117–147.
30. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
31. **Stearns, T., and D. Botstein.** 1988. Unlinked noncomplementation: isolation of new conditional-lethal mutations in each of the tubulin genes of *Saccharomyces cerevisiae*. Genetics **119**:249–260.
32. **Winston, F., D. T. Chaleff, B. Valent, and G. R. Fink.** 1984. Mutations affecting Ty-mediated expression of the *HIS4* gene of *Saccharomyces cerevisiae*. Genetics **107**:179–197.