Isolation of the Gene Encoding the *Saccharomyces cerevisiae* Centromere-Binding Protein CP1

RICHARD E. BAKER* AND DANIEL C. MASISON

Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, Massachusetts 01655

CP1 is a sequence-specific DNA-binding protein of the yeast *Saccharomyces cerevisiae* which recognizes the highly conserved DNA element I (CDEI) of yeast centromeres. We cloned and sequenced the gene encoding CP1. The gene codes for a protein of molecular weight 39,400. When expressed in *Escherichia coli*, the CP1 gene directed the synthesis of a CDEI-binding protein having the same gel mobility as purified yeast CP1. We have given the CP1 gene the genetic designation *CEP1* (centromere protein 1). *CEP1* was mapped and found to reside on chromosome X, 2.0 centimorgans from *SUP4*. Strains were constructed in which most of *CEP1* was deleted. Such strains lacked detectable CP1 activity and were viable; however, *CEP1* gene disruption resulted in a 35% increase in cell doubling time and a ninefold increase in the rate of mitotic chromosome loss. An unexpected consequence of CP1 gene disruption was methionine auxotrophy genetically linked to *cep1*. This result and the recent finding that CDEI sites in the *MET25* promoter are required to activate transcription (D. Thomas, H. Cherest, and Y. Surdin-Kerjan, J. Mol. Biol. 9:3292–3298, 1989) suggest that CP1 is both a kinetochore protein and a transcription factor.

CP1 is a sequence-specific DNA-binding protein of the yeast *Saccharomyces cerevisiae* which binds to both gene promoter regions and centromeres (6). The DNA recognition sequence for CP1 is the octanucleotide RTCACRTG (where R is a purine), which is found in all yeast centromeres and is named CDEI (centromere DNA element I) (24). The conservation of CDEI strongly implies a role for CP1 in the function of the centromere or kinetochore; however, CDEI sites are also found at many noncentromeric locations in the yeast genome (5, 6, 17). The existence of noncentromeric CDEI sites, together with the fact the CP1 is relatively abundant (greater than 500 molecules per cell), suggests that CP1 may play some general role in chromatin structure or in mediating DNA-protein interactions (5, 6).

Purified CP1 consists of a single polypeptide which appears to have a molecular weight of about 60,000 based on sodium dodecyl sulfate (SDS) gel analysis (5, 27). The quaternary structure of the native protein is not known. CP1 is unusually stable; it retains its specific DNA-binding activity after boiling and after elution and renaturation from SDS gels (5, 6). In its stability and DNA-binding properties, CP1 is similar to a protein in HeLa extracts which appears to be the adenovirus major late transcription factor (named MLTF or USF) (6, 12, 13, 42). Cai and Davis have purified a CDEI-binding protein from yeast cells which they call CBP-I (9). CBP-I has a molecular weight of 16,000 but has DNA-binding properties identical to those of CP1. These authors detected no CDEI-binding protein with a higher molecular weight, and they surmise the CBP-I could be derived from CP1, either through a specific posttranslational processing event or as a result of nonspecific protein degradation in their nuclear extracts (9).

CDEI is not essential for centromere function, but centromeres lacking an intact CDEI are impaired. Mitotically, chromosomes carrying CDEI mutations are lost at rates 2- to 60-fold higher than the corresponding chromosomes having wild-type centromeres (11, 15, 20, 23). Meiotically, chromo-

somal CDEI mutation or deletion has no detectable effect (20); however, centromere-containing plasmids (CEN plasmids) lacking CDEI exhibit random segregation at meiosis I (15, 20). The effects of CDEI mutations on CP1 binding and chromosome or CEN plasmid loss rates are quantitatively correlated (5, 9), directly implicating CP1 in mitotic centromere function; however, the exact action of CP1 remains obscure. At least one other protein is known to bind to centromeric DNA (22, 37). It binds to another conserved sequence, CDEIII, located approximately 100 base pairs (bp) away from CDEI. CDEIII is essential for proper centromere function (22, 23, 35, 37). Whether CP1 and the CDEIII-binding protein interact or influence each other's DNA binding is unknown.

This article describes the isolation of the gene encoding CP1 and the analysis of yeast strains lacking CP1. We show that the CP1 gene is not essential for growth, but that gene disruption leads to increased rates of chromosome and CEN plasmid loss. Furthermore, CP1 gene disruption has pleiotropic effects on cell metabolism seemingly unrelated to centromere function, and we will present evidence suggesting that CP1 is also a transcription factor.

MATERIALS AND METHODS

Strains and media. Strains used in this study are given in Table 1. JH15 was obtained from Jodi Hirschman of our department; other strains were constructed by us by standard genetic methods (36). All media were as described by Sherman et al. (43) except supplemented minimal medium, which contained 0.17% yeast nitrogen base (without amino acids), 0.1% ammonium sulfate, 2% glucose, with amino acids (40 µg/ml), adenine (20 µg/ml), and uracil (20 µg/ml) added as needed. Yeast transformations were performed by the lithium acetate procedure (26).

Plasmids. Plasmid pDR4-2 contains the 3.6-kilobase-pair (kbp) *Eco*RI fragment of the λ phage CP1-9.1.3 (see below) inserted into the *Eco*RI site of pGEM4 (Promega Corp., Madison, Wis.). Plasmid pRB85a-1 was derived from pDR4-2 by limited *Ssp*I digestion and religation in the presence of

---

* Corresponding author.

TABLE 1. Strains

Strain[a]	Genotype
381G	*MATa cry1 ade2-1 his4-580 lys2 trp1 tyr1 SUP4-a*
R11-4C	381G *MATa leu2 ura3 bar1-1 TYR1*
R11-4D	381G *MATα leu2 ura3 bar1-1 TYR1*
R15-3A	381G *MATα leu2 ura3 bar1-1(?) TYR1*
R11-4B	381G *MATα leu2 ura3 bar1-1 TYR1 cep1::URA3*
R15-3C	381G *MATa leu2 ura3 bar1-1 TYR1 cep1::URA3*
R15-3B	381G *MATa leu2 ura3 cep1::URA3*
JH15	381G *MATa leu2 ura3 bar1-1 TYR1*
	381G *MATα leu2 ura3 bar1-1 TYR1*
R11	381G *MATa leu2 ura3 bar1-1 TYR1 cep1::URA3*
	381G *MATα leu2 ura3 bar1-1 TYR1 CEP1*
R16	381G *MATa LEU2 ura3 BAR1 tyr1 cep1::URA3*
	381G *MATα leu2 ura3 bar1-1 TYR1 cep1::URA3*
R17	381G *MATa LEU2 ura3 BAR1 tyr1 cep1::URA3*
	381G *MATα leu2 ura3 bar1-1 TYR1 CEP1*
R18	381G *MATa LEU2 ura3 BAR1 tyr1 CEP1*
	381G *MATα leu2 ura3 bar1-1 TYR1 CEP1*
4795-408[b]	*MATa leu2 ade2 ade3 his7 trp1 can1 sap3*
J15-13C[c]	*MATα ade2-1 lys2-1 trp5-2 leu1-12 can1-100 met4-1 ura1-1 SUP4-o*
D2	R15-3B crossed with J15-13C

[a] All except the last three strains listed are congenic to strain 381G (21).

[b] Congenic to A364A; obtained from L. Hartwell, University of Washington.

[c] Kurjan and Hall (30); *SUP4-o* was scored by its ability to suppress *ade2-1* (ochre).

synthetic *Xho*I linkers. pRB85a-1 lacks the region of DNA between the *Ssp*I sites located at positions 409 and 1398 in the CP1 DNA sequence (Fig. 1). Plasmid pDR8 was obtained by digesting pRB85a-1 with *Xho*I and *Bgl*II, filling in the staggered ends with Klenow polymerase, and religating. This results in the removal of an additional 0.6 kbp of CP1 3'-flanking DNA and regenerates both *Xho*I and *Bgl*II sites. Plasmid pDR12 was derived from pDR8 by inserting a 1.1-kbp *Sal*I fragment containing the yeast *URA3* gene (40) into the unique pDR8 *Xho*I site. Plasmids pDR20 and pDR21 were derived from pUC(TRP/CEN3) and pUC(TRP/ΔCDEI), respectively (5), by inserting a 2.2-kbp DNA fragment containing the yeast *LEU2* gene (3) into the *Bam*HI site of their polylinkers.

The plasmids used for the expression of CP1 in *Escherichia coli* were derived from pTP201 (supplied by Anthony Poteete of our department), which in turn was derived from pTac12 (2) by inserting an *Nco*I linker at the *Pvu*II cloning site. pTP201 was cleaved with *Nco*I, the overhanging ends were filled in with Klenow polymerase, and the plasmid was religated in the presence of 8-bp *Pvu*II linkers. A 1.75-kbp *Hin*dIII-*Dra*I fragment containing the CP1 gene from pDR4-2 was then ligated into the new *Pvu*II site after Klenow fill-in of the *Hin*dIII staggered end. Plasmids pRB89-1 and pRB89-3 resulted from this ligation. pRB89-1 contains the CP1 open reading frame fused in frame at codon 9 with the ATG originating from the *Nco*I linker of pTP201, and pRB89-3 contains the CP1 segment in the opposite (antisense) orientation. The sequence of the noncoding strand of the fusion segment of pRB89-1 is given below, with the CP1 nucleotides and amino acids underlined and the linker sequences shown in lowercase letters:

```
----(tac promoter)----AGGAAACAGcccatg gca gAG CTT TCT----
                      Met Pro Glu Leu Ser----
```

All plasmids containing the *tac* promoter were propagated in *E. coli* W3110 (*lacI*q) (7). The predicted molecular weight of the *E. coli*-expressed CP1 is 38,800.

Cloning the CP1 gene. Purified CP1 protein (5) was subjected to trypsin digestion, and the partial amino acid sequence of two peptides (NCT14 and NCT11) was determined. This analysis was performed at the Harvard Microchemistry Facility with the method described by Aebersold et al. (1). The peptide sequences were used to design two nondegenerate oligonucleotide probes, NCT14 and NCT11:

NCT14: AspGlnGlyLeuLeuSerGlnGluSerAsnAspGlyAsnIleAspSerAlaLeuLeuSer
       5'-GATCAAGGTTTGTTGTCTCAAGAATCTAACGATGGTAACATTGATTCTGC-3'

NCT11: AlaIleThrProSerAsnGluGlyValLysProAsnThr[Gly/Ser]Leu
       5'-GCTATTACTCCATCTAACGAAGGTGTTAAGCCAAACAC-3'

In all but one case (AAG for Lys), we chose codons predicted from yeast codon usage frequencies supplied with the DNASTAR computer software package (DNASTAR, Inc., Madison, Wis.). The NCT14 and NCT11 oligonucleotides were synthesized, labeled with [γ-^{32}P]ATP by using T4 polynucleotide kinase, and used to screen a commercially available (Clontech, Palo Alto, Calif.) yeast genomic λgt11 library (*S. cerevisiae* X2180 [ATCC 26109]) by hybridization (25). Positive signals were obtained at a frequency of about 1 per 3,500 recombinant plaques. Two phages which hybridized to both probes were purified and found to contain overlapping yeast DNA inserts. One of them, λCP1-9.1.3, contained a 3.6-kbp *Eco*RI insert which was subcloned into both plasmid and M13 vectors.

The DNA sequence of the CP1 gene was determined by the dideoxy chain termination method with modified T7 DNA polymerase (44). The CP1 coding region was located immediately by using the NCT14 and NCT11 oligonucleotides as sequencing primers. Restriction sites discovered in these initial sequence runs were then used to map the location of the gene within the 3.6-kbp *Eco*RI fragment and to produce additional subclones. The CP1 DNA sequence presented in Fig. 1 was determined completely for both strands. A homology search of the National Biomedical Resource Facility protein data base (release no. 21) was performed against the deduced CP1 amino acid sequence. The search employed the Lipman-Pearson search algorithm (33) and was carried out by the PROSCAN computer program of DNASTAR.

Southwestern (protein-DNA) blots. *E. coli* containing the tac-CP1 expression plasmids pRB89-1 (sense) and pRB89-3 (antisense) were grown and induced with isopropyl-β-D-thiogalactoside (IPTG) as described by Amann et al. (2). After induction for 90 min, the cells from 1 to 2 ml of culture were sedimented by brief centrifugation and suspended in 75 to 150 μl of SDS sample buffer (63 mM Tris hydrochloride [pH 6.8], 2% SDS, 5% β-mercaptoethanol, 10% glycerol, 0.002% bromophenol blue. The ratio of cells to buffer was kept constant in order to obtain equivalent protein concentrations among samples; 0.6 OD_{600} units of cells were suspended per 75 μl of sample buffer. The samples were incubated in a boiling-water bath for 5 min, and then 5 μl (for Coomassie staining) or 10 μl (for blotting) of each was immediately loaded onto a 10 or 12% polyacrylamide (acrylamide-bisacrylamide, 37.5:1)-SDS gel prepared and run with the discontinuous buffer system of Laemmli (31). After electrophoresis, proteins were transferred to nitrocellulose membranes (Bio-Rad Laboratories) following the electroblotting procedure of Matsudaira (34).

After protein transfer, the membranes were blocked for 1 h at room temperature in 10 mM HEPES [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, pH 7.9]-1 mM dithiothreitol-1% nonfat dry milk and then incubated for 30

GGATCCTTCGGGACCACCATGTCTCACCCACCAAGGTCAAAAAAAATAAAATCGGACGAGAAAAGTATTGGGCAAACAACAATTATCCACGTGCTTATCA   100

CGTTTTTTTATTTAATGTAGAATGTAAACCATTATACCTTAAAGATAAACACTTTGTTAATTATTCGTTTTAGTTGGCGGTCATTTTCATTTCTGCTACA   200

                                                                              Hind III
AGTACCAACATCAAGTGCTTAAAAATATAATACGGTTTTCTACACTTTTATTAACGATGAACTCTCTGGCAAATAATAATAAGCTTTCTACTGAGGATGAG   300
                                                                M   N   S   L   A   N   N   N   K   L   S   T   E   D   E

GAAATCCATTCGGCAAGAAAAAGAGGCTATAACGAGGAGCAGAACTACAGCGAGGCCAGGAAAAAGCAAAGAGATCAAGGCTTGTTGTCACAAGAGAGCA   400
16   E   I   H   S   A   R   K   R   G   Y   N   E   E   Q   N   Y   S   E   A   R   K   K   Q   R   D   Q   G   L   L   S   Q   E   S   N
             Ssp I   NCT14
ATGACGGAAATATTGACTCTGCTCTGTTGAGCGAAGGGGCTACGCTAAAAGGGACTCAAAGTCAGTATGAAAGCGGACTGACGTCTAACAAAGATGAGAA   500
50   D   G   N   I   D   S   A   L   L   S   E   G   A   T   L   K   G   T   Q   S   Q   Y   E   S   G   L   T   S   N   K   D   E   K

AGGAAGTGATGACGAAGATGCGTCAGTGGCTGAGGCTGCTGTTGCCGCTACCGTCAATTATACAGACTTAATCCAGGGCCAGGAAGATAGTTCTGATGCC   600
83   G   S   D   D   E   D   A   S   V   A   E   A   A   V   A   A   T   V   N   Y   T   D   L   I   Q   G   Q   E   D   S   S   D   A

CACACATCTAATCAAACGAACGCGAATGGTGAGCACAAGGATTCCCTTAATGGGGAAAGAGCCATAACCCCCATCAAATGAAGGAGTCAAGCCAAACACCT   700
116 H   T   S   N   Q   T   N   A   N   G   E   H   K   D   S   L   N   G   E   R   A   I   T   P   S   N   E   G   V   K   P   N   T   S
                                                                               NCT11
CCCTGGAGGGAATGACGTCTTCACCAATGGAATCTACACAGCAATCTAAAAACGACATGCTCATACCATTGGCAGAACACGATCGTGGGCCAGAACACCA   800
150 L   E   G   M   T   S   S   P   M   E   S   T   Q   Q   S   K   N   D   M   L   I   P   L   A   E   H   D   R   G   P   E   H   Q

ACAGGATGATGAAGACAATGATGACGCTGACATTGATCTCAAAAAGGATATAAGCATGCAGCCGGGTCGTCGTGGAAGAAAACCTACTACTTTGGCCACA   900
183 Q   D   D   E   D   N   D   D   A   D   I   D   L   K   K   D   I   S   M   Q   P   G   R   R   G   R   K   P   T   T   L   A   T

ACAGACGAGTGGAAAAAGCAAAGAAAAGACTCCCATAAAGAAGTCGAAAGGCGTCGCCGCGAAAACATCAACACTGCAATCAACGTTTTAAGCGACCTCC   1000
216 T   D   E   W   K   K   Q   R   K   D   S   H   K   E   V   E   R   R   R   R   E   N   I   N   T   A   I   N   V   L   S   D   L   L
                                                Ssp I
TGCCCGTGAGAGAATCAAGTAAGGCAGCAATATTGGCGCGTGCCGCTGAATACATTCAAAAATTGAAAGAAACCGATGAAGCAAACATCGAAAAGTGGAC   1100
250 P   V   R   E   S   S   K   A   A   I   L   A   R   A   A   E   Y   I   Q   K   L   K   E   T   D   E   A   N   I   E   K   W   T

GTTACAAAAATTGCTTAGCGAGCAAAACGCATCGCAATTAGCCAGTGCAAATGAGAAACTGCAGGAAGAACTGGGAAATGCTTACAAAGAAATCGAGTAC   1200
283 L   Q   K   L   L   S   E   Q   N   A   S   Q   L   A   S   A   N   E   K   L   Q   E   E   L   G   N   A   Y   K   E   I   E   Y

ATGAAACGCGTTTTTAAGGAAGGAGGGAATAGAATACGAGGATATGCACACTCACAAGAAACAAGAGAACGAAAGAAAAAGCACTAGGAGCGATAATCCAC   1300
316 M   K   R   V   L   R   K   E   G   I   E   Y   E   D   M   H   T   H   K   K   Q   E   N   E   R   K   S   T   R   S   D   N   P   H
                                                                                               Ssp I
ATGAGGCTTGAGAGTTAAAAATAGATAGCTAAATGTATTTCGAGTCTCCCTATGTATCTGCGTCTCTGCCCCATAGCACCTCTGTTAAGTTTTTATCAAT   1400
350 E   A   *

ATT   1403
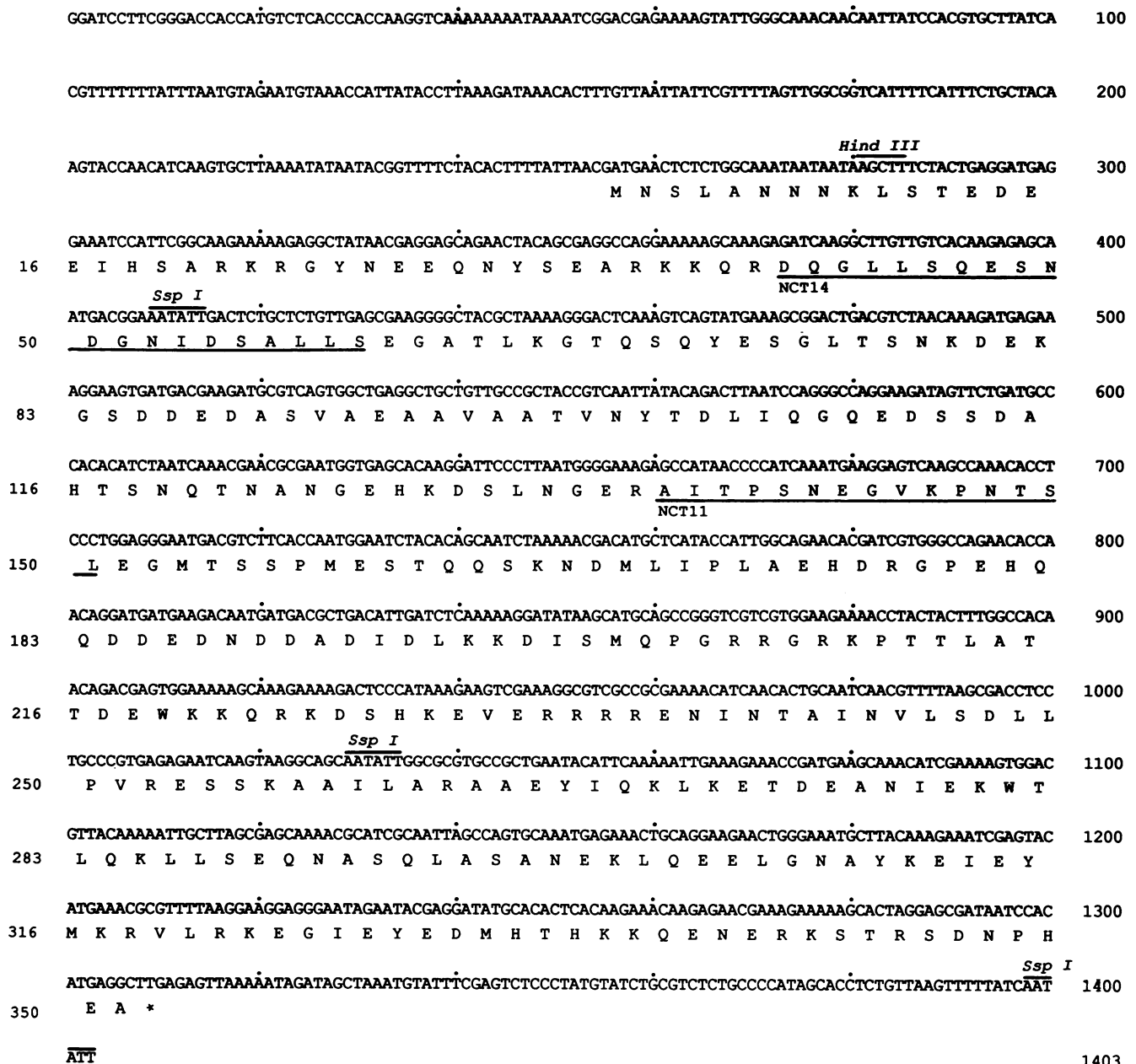
FIG. 1. DNA sequence of the gene encoding CP1. The DNA sequence is shown along with the protein translation of the 39-kDa open reading frame. The sequences of the NCT14 and NCT11 peptides are underlined. The *Ssp*I sites used to construct the *CEP1* gene disruption and the *Hin*dIII site used for the *tac* promoter fusion are overlined.

min in binding buffer (10 mM HEPES [pH 7.9], 50 mM NaCl, 10 mM MgCl$_2$, 0.1 mM EDTA, 1 mM dithiothreitol) containing ^{32}P-labeled CDEI DNA (7 ng/ml; 5 × 10^4 cpm/ng) and 5 μg of salmon sperm DNA per ml. Unbound probe was removed by two 7-min washes in binding buffer, and the membrane was blotted dry and used to expose X-ray film. In order to estimate protein molecular weights, prestained molecular weight standards (Bethesda Research Laboratories) were coelectrophoresed on the SDS gels and transferred to the membranes along with the sample proteins. Before autoradiography, the positions of the standards were marked with Glo-juice phosphorescent ink (International Biotechnologies, Inc.), allowing their direct identification on the autoradiograms.

The CDEI probe was a multimer (50 to 300 bp) of the synthetic double-stranded oligonucleotide shown below (CDEI site underlined):

5'-GATCCAAATAAGTCACATGATGATA-3'
GTTTATTCAGTGTACTACTATCTAG-5'
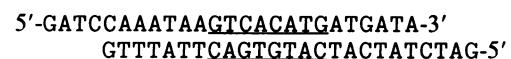
The oligonucleotide was phosphorylated, self-ligated, and labeled with [α-^{32}P]dCTP (800 Ci/mmol) by nick translation (46). CDEI competitor DNA was prepared by the same method, but it was of high molecular weight (>500 bp) and unlabeled.

Detection of CP1 in yeast extracts. Crude yeast extracts were prepared and gel shift DNA binding assays were performed as described previously (5). The CDEI probe

consisted of a hexamer of the synthetic CDEI oligonucleotide shown above, which was cloned into the BamHI site of pUC18, excised as a 205-bp EcoRI-HindIII fragment, and end-labeled with [α-^{32}P]dATP (800 Ci/mmol) by using Klenow polymerase. The tDNA probe was a 480-bp BglII fragment containing the yeast tRNA$_{UCG}^{Ser}$ gene (4) end-labeled with [α-^{32}P]dATP (800 Ci/mmol) by using Klenow polymerase (in the presence of unlabeled dGTP). Each binding reaction mix contained 10^4 cpm of probe (2 to 10 ng) and 6 μg of unlabeled poly[d(IC)]. Where indicated, binding reaction mixes received 5 μg of unlabeled poly[d(IC)] and 1.0 μg of the self-ligated CDEI competitor DNA described above.

CP1 gene disruption. We have designated the genetic locus of the CP1 gene CEP1 for centromere protein 1. CEP1 gene disruption was performed with the plasmid pDR12, in which most of the CP1 coding region (codons 53 to 351) and 0.6 kbp of 3'-flanking DNA were replaced by the yeast URA3 gene (see above). pDR12 was cleaved with EcoRI to excise the disrupted CP1 gene (cep1::URA3) from the vector and then used to transform the diploid strain JH15. Homologous recombination stimulated by the free ends of the transforming DNA results in replacement of the wild-type CP1 gene with the cep1::URA3 disruption allele (41). Ura$^+$ transformants were analyzed by Southern blotting, and most were found to contain one copy of the disrupted CP1 gene and one copy of the wild-type gene. One such transformant, R11, was sporulated, and the tetrads were dissected. The URA3 marker segregated 2+:2−, and Southern blotting showed that the Ura$^+$ spores gave rise to restriction fragments diagnostic of the gene disruption and Ura$^-$ spores gave rise to fragments diagnostic of the wild-type gene; i.e., the physical and genetic markers cosegregated.

Chromosome and plasmid loss assays. Chromosome III loss rates in the diploid strains R16, R17, and R18 were determined by quantitative mating (39). The haploid tester stain used was 4795-408 (MATa). Colonies of the diploids were grown on plates lacking leucine to require maintenance of the MATaLEU2 chromosome. For each determination, a single colony was suspended in water, sonicated briefly, and used to inoculate a 100-ml rich broth (YEPD) culture. A portion of the cell suspension was also diluted and plated on YEPD plates to determine the starting cell number. The YEPD liquid cultures were grown at 30°C to mid-log phase (16 to 18h), and the cell density was determined. In duplicate, 2 × 10^6 diploid cells were mixed with 2 × 10^6 haploid tester cells, also from a mid-log-phase culture, and filtered onto a 25-mm Millipore filter. Filters were incubated cell side up on YEPD plates at 30°C for 4.5 h to allow mating. The mating mixtures were then resuspended from the filters in 5.0 ml of water, diluted, and spread on plates lacking histidine and lysine to select for growth of zygotes. In each experiment, haploid strain R11-4D (MATα) was mated with strain 4795-408 under the same conditions in order to establish the mating efficiency. After 48 h, colonies were counted and then replicated onto plates lacking leucine to test for the presence of the LEU2 marker originally present on the MATa chromosome. Leu$^-$ colonies were assumed to have arisen from cells which had become competent to mate due to loss of the MATaLEU2 chromosome during the nonselective growth of the culture. The frequency of Leu$^-$ zygotes (corrected for mating efficiency) is therefore an indirect measure of chromosome loss frequency. Chromosome loss rates were calculated from the following formula (18): loss rate = (0.43)(loss frequency)/(log N − log N_0), where N and N_0 are the number of cells in the population at the end and start of nonselective (YEPD) growth, respectively. Control

experiments showed that the cep1::URA3 allele does not affect the mating efficiency of haploid strains of either mating type (data not shown). We do not know whether chromosome III monosomy in our diploid strains leads to any growth defect. If it does, our results would be skewed; mating-competent cells (loss events) would be underrepresented in the total cell population, leading to an underestimation of the loss rate.

Plasmid loss determinations were carried out essentially as described before (5) except the plasmid selectable marker was LEU2. Individual transformant colonies were grown selectively, either in liquid or on plates, to maintain the CEN plasmid. After selective growth, cell suspensions were used to inoculate liquid YEPD cultures and at the same time diluted and spread on YEPD plates to obtain single colonies (t_0). After 18 to 20 h of growth at 30°C, the liquid YEPD cultures were diluted and plated for single colonies on YEPD plates (t_n). After colonies had formed (2 days), the t_0 and t_n plates were replicated onto supplemented minimal plates with or without leucine, and after 24 h the percentage of Leu$^+$ colonies was determined. The number of cell doublings was calculated by determining the total cell number at the start and end of nonselective (YEPD) growth. Plasmid loss rates were then calculated from the following equation (29): $L_n/L_0 = (1 - R)^n$, where L_n and L_0 are the percentage of Leu$^+$ colonies after nonselective and selective growth, respectively, n is the number of doublings during nonselective growth, and R is the plasmid loss rate (loss events per cell per generation).

RESULTS

Isolation of the gene encoding CP1. Purified CP1 was cleaved with trypsin, and the partial amino acid sequences of two tryptic peptides, NCT14 and NCT11, were determined. The sequences were used to design two oligonucleotide probes, a 50-mer from NCT14 and a 38-mer from NCT11, which were synthesized, labeled, and used to screen a λgt11 yeast genomic library by hybridization (details are given in Materials and Methods). Two different phages were obtained which hybridized to both probes, and the 3.6-kbp EcoRI insert from one of them was subcloned for DNA sequence analysis. The complete DNA sequence of the CP1 gene is given in Fig. 1 along with the deduced CP1 amino acid sequence. The gene codes for a protein of molecular weight 39,400. A homology search of the National Biomedical Resource Facility protein sequence data base against the predicted CP1 amino acid sequence revealed no striking similarity to any protein in the data base.

The molecular weight of CP1 determined by SDS gel electrophoresis is 58,300 (5), significantly larger than that of the polypeptide encoded by the putative CP1 gene. Of the several possible explanations for this discrepancy, we though it most likely that the CP1 protein contains posttranslational modifications which affect its mobility in SDS gels. To test this and at the same time prove that the open reading frame we identified actually codes for CP1, we expressed the gene in E. coli. A hybrid transcription unit was constructed by fusing the CP1 open reading frame at the naturally occurring HindIII site at CP1 codon 9 to the E. coli tac promoter (2). An initiator ATG was provided by the tac promoter fragment, and the molecular weight of the expected translation product was 38,800.

Proteins produced by E. coli strains carrying tac-CP1 expression constructs were analyzed by Southwestern blotting. Proteins were separated on SDS gels and transferred to
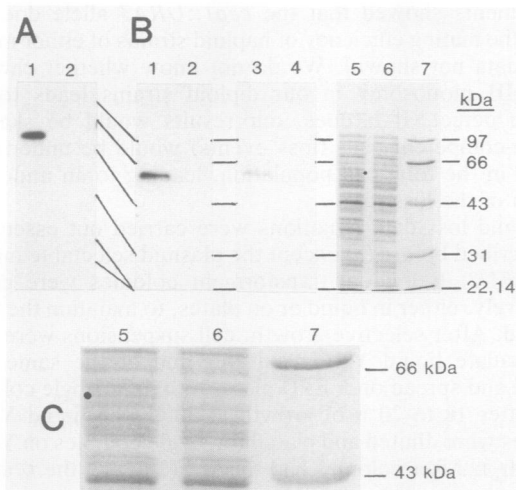
FIG. 2. Expression of CP1 in *E. coli*. Proteins from bacteria containing the *tac*-CP1 plasmid pRB89-1 (CP1 open reading frame in the sense orientation) or pRB89-3 (CP1 open reading frame in the antisense orientation) were analyzed for CDEI-binding activity by Southwestern blotting as described in Materials and Methods. (A) Proteins from a pRB89-1-bearing strain induced (lane 1) or uninduced (lane 2) by IPTG were analyzed on a 12% polyacrylamide resolving gel. (B) Proteins from IPTG-induced strains bearing pRB89-1 (lanes 1, 3, and 5) or pRB89-3 (lanes 2, 4, and 6) were analyzed on a 10% gel. After blotting, the membrane was cut, and one half was incubated with probe in the presence of a 25-fold excess of unlabeled CDEI DNA (lanes 3 and 4) or an equivalent amount of poly[d(IC)] (lanes 1 and 2). Lanes 5 to 7 show the results of Coomassie blue staining. Lane 7 contains protein molecular weight standards (BioRad). (C) Enlargement of the 40 to 70-kDa region of the Coomassie-stained gel. The putative CP1 translation product is indicated by dots (panels B and C, lanes 5).

nitrocellulose membranes, which were then incubated with a ^{32}P-labeled synthetic DNA containing CDEI, i.e., the recognition site for CP1. Figure 2A shows the results for bacteria grown in the presence or absence of IPTG, which is required to induce transcription of the *tac* promoter. Upon induction, a protein band was detected which bound the CDEI probe (Fig. 2A, lane 1). The binding activity was not present in protein extracts prepared from the same *E. coli* strain grown in the absence of IPTG (Fig. 2A, lane 2), nor was it observed in induced strains carrying a *tac*-CP1 construct in which the CP1 segment was inserted in the opposite (antisense) orientation (Fig. 2B, lane 2). Binding to the probe was significantly reduced in the presence of excess unlabeled CDEI-containing DNA (Fig. 2B, lane 3), but not in the presence of poly[d(IC)] or salmon sperm DNA.

Based on its mobility relative to prestained electrophoresis standards, the *E. coli*-produced CDEI-binding protein had a molecular mass of 55 to 60 kilodaltons (kDa). Close examination of Coomassie blue-stained gels revealed that a new protein band was present in this region of the gel (Fig. 2B and C, lanes 5); its apparent molecular-weight was 58,300, based on the Coomassie blue-stained standards. While not prominent, the appearance of this protein band correlated exactly with the synthesis of the binding activity, i.e., it was present only after IPTG induction and in strains carrying the sense *tac*-CP1 construct. The *E. coli*-synthesized protein, then, had the same mobility on SDS gels as purified yeast CP1. We conclude that the gene we have isolated does indeed code for the CDEI-binding protein CP1. Since we are unaware of any posttranslational modifications

TABLE 2. Tetrad analysis

Diploid	Interval	No. of asci[a]			Map distance[b] (centimorgans)
		PD	NPD	T	
D2	*cep1::URA3-SUP4*	72	0	3	2.0
R17	*cep1::URA3-met*	53	0	0	0.0
	cep1::URA3-tyr1	13	7	32	NL

[a] PD, Parental ditype; NPD, nonparental ditype; T, tetratype.
[b] Map distance = {[T + 6(NPD)]/[2(PD + NPD + T)]} × 100 (36). NL, No linkage.

common to *S. cerevisiae* and *E. coli* which would explain such a large discrepancy between the predicted molecular weight and that which was observed, we also conclude that the CP1 polypeptide migrates aberrantly in SDS gels.

Construction of a yeast strain with a disrupted CP1 gene. As a first step towards elucidating the function of CP1 in the cell, we constructed yeast strains in which the gene encoding CP1 was disrupted. Starting with a plasmid containing the cloned CP1 gene, most of the coding region (codons 53 to 351) along with 3'-flanking DNA was excised and replaced with a DNA fragment carrying the yeast *URA3* gene. The disrupted gene, along with its remaining flanking regions, was then excised from the plasmid and used to transform a diploid homozygous *ura3* strain. Ura$^+$ transformants were obtained in which the disrupted gene construct had replaced the resident CP1 gene on one of the two chromosomes. One such strain, R11, was analyzed by Southern blotting and found to have the expected CP1 gene configuration, i.e., restriction digests yielded one fragment derived from the parental gene and a second fragment derived from the disrupted gene (data not shown). When R11 was sporulated and the asci were dissected, most asci gave rise to three or four viable spores, indicating that the CP1 gene is not essential. The *URA3* marker, physically linked to the disrupted CP1 gene, segregated 2+:2−. We designated the genetic locus of the CP1 gene *CEP1*, for centromere protein 1.

The *CEP1* locus was mapped, first by chromosome blotting and subsequently by genetic linkage analysis. Hybridizing a Southern blot of electrophoretically separated yeast chromosomes with a *CEP1* probe allowed us to assign *CEP1* to chromosome X (data not shown). When genetic linkage to other chromosome X markers was tested, we found that *CEP1* mapped very close to *SUP4*. Table 2 gives the results of tetrad analysis. The genetic distance between *CEP1* and *SUP4* was 2.0 centimorgans.

CDEI-binding activity in *CEP1*$^+$ and *cep1::URA3* strains. Extracts were prepared from haploid segregants of R11 and assayed for CDEI-binding activity by gel shift assay. When a *CEP1*$^+$ extract (R11-4C) was incubated with a CDEI probe, a predominant CP1-DNA complex was formed as well as lesser amounts of two or more complexes of lower mobility (Fig. 3, lane 3). The lower-mobility complexes are probably oligomers of CP1. All of these complexes were also formed by purified CP1 (lane 2), and their formation was greater than 95% reduced in the presence of excess unlabeled CDEI-containing DNA (lanes 5 and 6; quantitation by densitometry). (The complexes observed in lane 2 migrating beneath the major CP1-DNA complex were due to partial degradation of CP1, which occurred during its purification). An extract prepared from a *cep1::URA3* strain (R11-4B) failed to form any of the CP1-DNA complexes (lane 4), although a faint, smeared signal was detected in the region of the gel at the position of the low-mobility complexes. We do
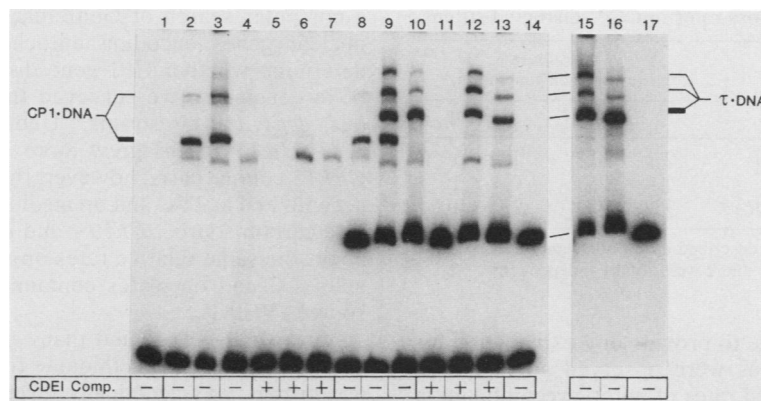
FIG. 3. CDEI-binding activity in *CEP1*[+] and *cep1::URA3* strains. Crude whole-cell extracts of the *CEP1*[+] strain R11-4C (lanes 3, 6, 9, 12, 15) and the *cep1::URA3* strain R11-4B (lanes 4, 7, 10, 13, 16) were prepared and assayed for specific DNA-binding activity by gel shift assay as described in Materials and Methods. The reaction mixes in lanes 2, 5, 8, and 11 contain purified CP1 (5). The labeled probes were: CDEI only, lanes 1 to 7; both CDEI and tDNA, lanes 8 to 14; tDNA only, lanes 15 to 17. Where indicated below the figure, an approximate 500-fold excess of unlabeled CDEI competitor (comp.) DNA was added to the binding reaction mixes. For the crude extract assays, equal amounts of protein (25 μg) were used per reaction. Lanes 1, 14, and 17 contain no added protein.

not believe that this weak signal is due to residual CP1, because in dozens of gel shift assays with partially pure or pure CP1, we never observed low-mobility CP1 complexes in the absence of a major CP1-DNA band. A distinct band which ran below the main CP1-DNA complex was observed with both the R11-4B and R11-4C extracts, but this band was due to a nonspecific binding activity since its intensity was not reduced at all by the CDEI competitor DNA.

To ensure that our failure to detect CP1 activity in the R11-4B extract was not simply due to improper handling of the extract or to some general effect of the *CEP1* disruption on DNA-binding proteins (e.g., increased levels of protease), we simultaneously assayed for another sequence-specific DNA-binding protein, transcription factor τ. Factor τ is an RNA polymerase III transcription factor which binds to the internal promoter of tRNA genes (10). Factor τ-binding activity was readily detected in both R11-4C and R11-4B crude extracts using a tDNA^Ser probe (Fig. 3, lanes 15 and 16). When both probes were added together, both CP1 and τ complexes were observed with the R11-4C extract (lane 9), but only τ complexes were observed with the R11-4B extract (lane 10). Only the CP1 complexes were reduced in the presence of CDEI competitor DNA (lanes 11 to 13). We conclude that the *cep1::URA3* disruption results in the loss of functional CP1 protein from the cell.

Phenotypic effects of *CEP1* gene disruption. Haploid strains carrying the *cep1::URA3* gene disruption were relatively healthy; however, we noticed immediately that *cep1::URA3* disruptants grew more slowly than isogenic *CEP1*[+] strains. The doubling times for several strains, both haploid and diploid, growing in rich medium at 30°C are given in Table 3. The *cep1::URA3* gene disruption led to a 35% lengthening in the population doubling time, from about 95 to about 130 min. The growth defect was recessive; *cep1::URA3/+* heterozygous diploids grew at the same rate as *CEP1*[+] haploids or homozygous *CEP1*[+] diploids.

Since CP1 binding seems to be required for optimal mitotic function of the centromere (5), a yeast strain lacking CP1 should exhibit increased rates of chromosome loss. Chromosome III loss rates in *cep1::URA3* strains were determined by a quantitative mating assay (14). The assay is based on the fact that a *MATa/MATα* diploid cell, a nonmater, can become competent to mate if it loses one of its chromosome III homologs. (The mating type locus, *MAT*, is located on

the right arm of chromosome III.) Chromosome III loss results in a cell's becoming phenotypically **a** or α, depending on which homolog is lost, and able to mate with a haploid cell of the opposite mating type. An **a**/α diploid can also become competent to mate through mitotic recombination or gene conversion at the *MAT* locus; however, these events can be distinguished from chromosome loss experimentally by scoring for the concomitant loss of a genetic marker on the opposite arm of chromosome III.

Table 4 gives the results of two separate experiments in which we measured rates of chromosome III loss from the diploid strains R16 (*cep1::URA3/cep1::URA3*), R17 (*CEP1*[+]/*cep1::URA3*), and R18 (*CEP1*[+]/*CEP1*[+]). The rate of chromosome III loss from the wild-type diploid (R18) was 1×10^{-5} to 2×10^{-5} events per cell per division. This value agrees with previous measurements of chromosome III loss (14, 19). When homozygous, the *cep1::URA3* disruption caused a ninefold increase in the rate of chromosome III loss. As with the slow-growth phenotype, the gene disruption was recessive in its effect on chromosome loss. In fact, the rate of loss in the *cep1::URA3/+* heterozygote (R17) seemed to be even lower than that of the wild-type diploid (R18). This twofold difference was observed consistently,

TABLE 3. Doubling times of *CEP1*-disrupted strains[a]

Strain	*CEP1* genotype	Doubling time[b] (min)
Haploids		
R11-4C	*CEP1*[+]	98 ± 4 (2)
R15-3A	*CEP1*[+]	95
R11-4B	*cep1::URA3*	128 ± 4 (2)
R15-3C	*cep1::URA3*	133
Diploids		
R18	*CEP1*[+]/*CEP1*[+]	98
R17	*CEP1*[+]/*cep1::URA3*	94
R16	*cep1::URA3/cep1::URA3*	127

[a] Cells from late-log-phase cultures were diluted into fresh YEPD medium and grown at 30°C. Portions were removed at various times, and after brief sonication to separate cells, the cell number was determined with a Coulter counter.
[b] Doubling times were determined from semilog plots, fitting the curves by the method of least squares. When more than one determination was made (number shown in parentheses), the result is reported as the mean ± the total deviation.

TABLE 4. Chromosome III loss rates in *CEP1*-disrupted strains

Strain	*CEP1* genotype	Loss rate[a] (events/cell/generation)	
		Exp 1	Exp 2
R18	*CEP1⁺/CEP1⁺*	2.2×10^{-5}	1.1×10^{-5}
R17	*CEP1⁺/cep1::URA3*	1.4×10^{-5}	0.4×10^{-5}
R16	*cep1::URA3/cep1::URA3*	1.9×10^{-4}	1.0×10^{-4}

[a] Loss rates were determined as described in Materials and Methods. The strain stocks used in experiments 1 and 2 were totally independent.

but at present we are unable to provide any explanation for it. No significant differences were observed among these three strains in the combined rates of mitotic recombination and gene conversion (data not shown).

Rates of minichromosome (i.e., CEN plasmid) loss were also determined for wild-type and *CEP1* disruption strains. In this case, it was possible to directly compare the *cis* effect of a CDEI mutation with the *trans* effect of disrupting *CEP1*. If the only involvement of CP1 in mitotic chromosome segregation is exerted through its binding to CDEI at the centromere, then the effect of deleting CDEI from the centromere in a normal cell should be quantitatively similar to the effect on a wild-type centromere when CP1 is absent from the cell. Loss rates were measured for two CEN plasmids, pDR20 and pDR21, in *CEP1⁺* and *cep1::URA3* haploid strains. The plasmids are identical except that pDR20 contains a wild-type CEN3 segment, whereas the CEN segment in pDR21 lacks CDEI owing to a 14-bp deletion. The plasmids also contain the selectable markers *TRP1* and *LEU2* and the replicator sequence (origin) *ARS1*.

Previous experiments (5) had shown that the CDEI deletion in pDR21 leads to an increased rate of plasmid loss, but actual loss rates were not determined. Table 5 shows that pDR20 was lost from *CEP1⁺* cells at a rate of 1.6% per division; the loss rate of pDR21 was increased 2.5-fold to 3.9% per division. The rate of loss of the wild-type CEN plasmid from the *cep1::URA3* strain was 3.2% per division, not significantly (or statistically) different from the rate of loss of the CDEI-deleted plasmid from the wild-type cell. Furthermore, the effects of the mutations were not additive; the mutant CEN plasmid was lost from the *cep1::URA3* strain at a rate of 3.6% per division. We conclude that the involvement of CDEI in mitotic centromere function is mediated through the *CEP1* gene product, CP1.

Methionine auxotrophy. Several researchers have noted that CDEI sites exist in many noncentromeric locations throughout the yeast genome. Several genes contain CDEI sites in their 5′-flanking regions, where they could conceivably serve as transcriptional activation sequences (UASs), among them *MET2*, *MET25*, *ADE3*, *GPA2*, and *SRA3* (from

TABLE 5. CEN plasmid loss rate

Strain	Plasmid	Mean % loss/cell/ generation[a] ± SEM
R11-4C (*CEP1⁺*)	pDR20 (CEN)	1.6 ± 0.2 (9)
	pDR21 (ΔCDEI)	3.9 ± 0.8 (8)
R11-4B (*cep1::URA3*)	pDR20 (CEN)	3.2 ± 0.3 (9)
	pDR21 (ΔCDEI)	3.6 ± 0.3 (8)

[a] The data from three separate experiments were pooled; the number of independent determinations for each strain-plasmid combination is given in parentheses. By the F test (32), the only statistically significant difference among the means ($P > 0.95$) was the loss rate of plasmid pDR20 in strain R11-4C, which was different from each of the other three.

a computer search of GenBank, release no. 58) and three nuclear genes encoding mitochondrial proteins (17). To determine whether CP1 gene disruption caused any metabolic deficiency, we surveyed the growth requirements of our *cep1::URA3* strains. Under all conditions tested, *cep1::URA3* strains grew more slowly than their isogenic *CEP1⁺* counterparts; however, the growth disadvantage was accentuated at 22°C and on medium containing lactate as the sole carbon source. *CEP1⁺* and *cep* mutant strains grew at about the same relative rates on rich medium (YEPD) at 30 and 34°C and on plates containing glycerol as the carbon source (YEPG).

Surprisingly, we found that *cep1::URA3* strains required exogenously added methionine for growth. The methionine auxotrophy was genetically linked to the *cep1::URA3* disruption (Table 2). In 53 tetrads dissected from a *cep1::URA3/+* diploid, no Ura⁺ Met⁺ or Ura⁻ Met⁻ spores were obtained, meaning that the lesion which gives rise to the methionine auxotrophy was not genetically separable from the *cep1::URA3* gene disruption. (The unlinked *tyr1* marker [chromosome II] showed the 1:1:4 segregation pattern expected.) Formally, the genetic distance between *cep1::URA3* and *met* is less than 0.9 centimorgan (one tetratype in 54 meioses). We interpret this result to mean that the *cep1::URA3* gene disruption in fact causes the methionine auxotrophy. None of the 21 genetically defined *MET* genes has been mapped to this region of chromosome X, although several have not yet been mapped (28).

DISCUSSION

We have cloned the gene encoding the yeast centromere-binding protein CP1. The molecular weight of CP1 predicted from its DNA sequence is 39,400, considerably less than the apparent molecular weight of the purified protein, which was 58,000 as determined by SDS gel analysis. However, when expressed in *E. coli*, the cloned CP1 gene directs the synthesis of a protein whose mobility in SDS gels is indistinguishable from that of native yeast CP1 and which specifically binds to DNA containing CDEI, the DNA recognition sequence of CP1. Furthermore, the deduced CP1 protein sequence contains exact matches to the partial amino acid sequences that were determined for two CP1 tryptic peptides. These results along with the genetic evidence discussed below leave little doubt that the gene isolated does indeed encode CP1. Evidently, CP1 migrates aberrantly in SDS gels. The predicted CP1 protein sequence contains 20% acidic residues (glutamate, aspartate), which results in the protein's having a predicted net charge of −20 at pH 7. Whether it is this strongly acidic character or perhaps an unusual secondary structure which leads to the anomalous SDS gel mobility is unknown. A similar situation exists with the yeast protein ARS-binding factor 1 (ABF1) (16). The ABF1 gene encodes a polypeptide of molecular weight 81,759, yet purified ABF1 migrates on SDS gels with an apparent molecular weight of 135,000. Like CP1, ABF1 has a high aspartic acid content (11.4%) and high net negative charge at neutral pH (−6.0). The yeast transcription factors HSTF (heat shock transcription factor) and RAP1 also migrate slowly in SDS gels (references in reference 16).

Cloning the CP1 gene, which we designated *CEP1*, enabled us to obtain yeast strains in which this gene is disrupted. Haploid cells containing a disrupted *CEP1* gene are viable, indicating the *CEP1* is nonessential. However, *cep1* strains grow more slowly than their isogenic *CEP1⁺* counterparts, exhibit increased rates of mitotic chromosome

and CEN plasmid loss, and are methionine auxotrophs. To construct the *cep1::URA3* disruption allele (strain R11-4B), 900 bp of coding region and about 700 bp of 3'-flanking DNA were deleted. A possible explanation for the pleiotropic effects of this mutation is that the deletion of downstream sequences has disrupted a neighboring gene. Three lines of evidence (data not shown) argue against this possibility. First, strains containing an insertion of the yeast *LEU2* gene 700 bp downstream of *CEP1* (the *Bgl*II site) grew normally and were methionine prototrophs. Second, a haploid strain containing a *cep1::URA3* disruption which removes only coding region and 90 bp of downstream sequences (to the *Ssp*I site at position 1398, Fig. 1) had a phenotype indistinguishable from that of R11-4B, i.e., slow growth, increased CEN plasmid loss, and methionine auxotrophy. Third, the methionine auxotrophy linked to the R11-4B gene disruption was complemented by a CEN plasmid containing the *CEP1* gene having only 110 bp of 3'-flanking DNA.

The observed increase in culture doubling time of *cep1* strains was about 35%, while chromosome loss rate was increased ninefold. Mathematical modeling (not shown) reveals that to cause a 35% increase in population doubling time, chromosome loss (or any event leading to cell inviability) would have to occur in about 10 out of 25 mitoses, or 40% per cell per division. The increased rate of chromosome loss, from about 10^{-5} loss events per cell per division in a wild-type strain to 10^{-4} loss events per cell per division in a *cep1* strain, is not nearly sufficient to account for the decreased rate of culture growth. Also, we did not observe an accumulation of inviable cells in cultures of *cep1* strains. Although other explanations are possible, we suggest that *cep1* cells actually have longer cell cycle times, perhaps implying that the role of CP1 in cell metabolism is not confined to mitosis.

Gel shift DNA-binding assays performed with extracts from haploid *cep1::URA3* cells revealed no detectable CP1. This was not surprising, since the results of genomic Southern blots indicated that the *CEP1* gene is unique. But is CP1 the only CDEI-binding protein in the cell? It is not totally inconceivable, for instance, that CP1 carries out a function related to noncentromere CDEI sites while an as yet unrecognized, centromere-specific CDEI-binding protein is responsible for CDEI-related centromere function(s). Our DNA-binding assay results argue against such a scenario. No clear CDEI-specific binding activity was detected at all in *cep1* cells. A weak, indistinct gel shift signal was observed, but it constituted at most 6% (quantitation not shown) of the CDEI-binding activity present in *CEP1*$^+$ cells. We favor the view that CP1 is the only CDEI-binding protein in *S. cerevisiae*; however, our data do not rule out the existence of others.

Point mutations or deletion of CDEI from the centromeres of plasmid minichromosomes results in small but significant decreases in their mitotic stability (5, 23, 38). Our results here show that the increase in plasmid loss rate due to CDEI deletion is about 2.5-fold. Hegemann et al. tested the effect of several CDEI point mutations on the mitotic stability of 150-kbp chromosome fragments and found 2- to 10-fold-elevated loss rates depending on the mutation (23). The effect of deleting CDEI from CEN3 is a 20- to 60-fold increase in the rate of mitotic chromosome loss (15, 20). We found that a cell lacking CP1 loses chromosome III at a rate ninefold higher than that of a wild-type cell. This observed increase in mitotic loss rate is somewhat less than might be expected from the results of the CDEI deletion analyses just

described but is still consistent with our assertion that the mitotic function of CDEI is mediated through CP1 (5).

The *trans* effect of removing CP1 from the cell and the *cis* effect of deleting CDEI from the centromere were directly compared by measuring the loss rates of CEN plasmids from wild-type and *cep1::URA3* strains. The rate of loss of a wild-type CEN plasmid from a *cep1* mutant cell was not significantly different from the rate of loss of a CDEI-deleted CEN plasmid from a *CEP1*$^+$ cell. Furthermore, the *cis* and *trans* effects were not additive in that the rate of loss of the CDEI-deleted plasmid in the *cep1* mutant strain was not further increased. These results indicate that the impaired function of CDEI-mutated centromeres can be attributed to the lack of CP1 interaction. It follows then that CP1 is directly involved in centromere function and that no other CDEI-binding protein exists which can fulfill this role.

CDEI sites are found at several noncentromere locations in the yeast genome, in some instances in gene promoter regions (5, 6, 17). In fact, in their original paper, Bram and Kornberg speculated that in some contexts CP1 might function as a transcription factor (6). Interestingly, CP1 has a DNA-binding specificity that is seemingly identical to that of the human transcription factor MLTF. However, partially purified CP1 cannot substitute for MLTF in a reconstituted mammalian in vitro transcription system (13). Recently, Thomas et al. discovered that CDEI sites upstream of the yeast *MET25* gene (coding for *O*-acetylhomoserine sulfhydrylase) are required for transcriptional activation (45). When the CDEI sequences are deleted, enzyme activity is decreased by over 90% and a corresponding decrease is observed in the steady-state levels of *MET25* mRNA (45). Strains with *met25* mutations require exogenously added methionine for growth. It seems highly likely that the methionine auxotrophy of our *cep1* mutant strains is caused at least in part by an inability to transcribe *MET25*. CDEI sites are also found in the promoter regions of *MET2*, *MET3*, *SAM1*, and *SAM2*, genes which are coregulated with *MET25* (45). This implicates CP1 in the transcriptional control of several genes involved in methionine metabolism, and it is possible that it is the disruption of these interrelated pathways which leads to the growth defect of *cep1* strains even when growing on rich medium.

As yet we have no direct evidence that CP1 functions as a transcription factor. We are in the process of analyzing *MET25* transcription in our *cep1::URA3* strains, and the results will tell whether or not CP1 is required for *MET25* expression. If CP1 is a transcription factor, we would like to know how it performs its dual role. Does it carry out exactly the same function at the centromere as it does at a gene promoter? Does it interact with the same or different proteins at its various sites of action?

The inferred multifunctional nature of CP1 is reminiscent of another yeast DNA-binding protein, general regulatory factor I (GRFI). GRFI (also known at RAP1) binds to a sequence element found in gene promoters (*MAT*α1 and *MAT*α2, *TEF2*), in the silencer regions of the *HML* and *HMR* mating type loci, and in yeast telomeres, thereby suggesting roles for GRFI in transcriptional activation, transcriptional repression, and chromosome structure (8). Both CP1 and GRFI are relatively abundant (5, 6, 8). Buchman et al. have suggested that GRFI and CP1 are members of a class of chromatin structural components which act in conjunction with other factors to build the chromatin structures necessary for both gene expression and chromosome maintenance (8). Our results with CP1 are not inconsistent with that view. More specifically, we are intrigued by the possi-

bility that CP1 and/or GRFI might bind to their cognate sites in DNA and then direct nuclear matrix attachment or determine intranuclear location. Cloning *CEP1* has allowed us to begin the genetic analysis of CP1 function. Hopefully, the detailed study of *cep1* mutants will enable us to further elucidate the role(s) of CP1 in cellular metabolism.

ACKNOWLEDGMENTS

LITERATURE CITED

1. Aebersold, R. H., J. Leavitt, R. A. Saavedra, L. E. Hood, and S. B. H. Kent. 1987. Internal amino acid sequence analysis of proteins separated by one- or two-dimensional gel electrophoresis after *in situ* protease digestion on nitrocellulose. Proc. Natl. Acad. Sci. USA **84**:6970–6974.

2. Amann, E., J. Brosius, and M. Ptashne. 1983. Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. Gene **25**:167–178.

3. Andreadis, A., Y. Hsu, G. B. Kohlhaw, and P. Schimmel. 1982. Nucleotide sequence of yeast *LEU2* shows 5'-noncoding region has sequences cognate to leucine. Cell **31**:319–325.

4. Baker, R. E., O. Gabrielsen, and B. D. Hall. 1986. Effects of tRNA^{Tyr} point mutations on the binding of yeast RNA polymerase III transcription factor C. J. Biol. Chem. **261**:5275–5282.

5. Baker, R. E., M. Fitzgerald-Hayes, and T. C. O'Brien. 1989. Purification of the yeast centromere-binding protein CP1 and a mutational analysis of its binding site. J. Biol. Chem. **264**:10843–10850.

6. Bram, R. J., and R. D. Kornberg. 1987. Isolation of a *Saccharomyces cerevisiae* centromere DNA-binding protein, its human homolog, and its possible role as a transcription factor. Mol. Cell. Biol. **7**:403–409.

7. Brent, R., and M. Ptashne. 1981. Mechanism of action of the *lexA* gene product. Proc. Natl. Acad. Sci. USA **78**:4204–4208.

8. Buchman, A. R., W. J. Kimmerly, J. Rine, and R. D. Kornberg. 1988. Two DNA-binding factors recognize specific sequences at silencers, upstream activating sequences, autonomously replicating sequences, and telomeres in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8**:210–225.

9. Cai, M., and R. W. Davis. 1989. Purification of a yeast centromere-binding protein that is able to distinguish single-base-pair mutations in its recognition site. Mol. Cell. Biol. **9**:2544–2550.

10. Camier, S., O. Gabrielsen, R. Baker, and A. Sentenac. 1985. A split binding site for transcription factor τ on the tRNA^{Glu} gene. EMBO J. **4**:491–500.

11. Carbon, J., and L. Clarke. 1984. Structural and functional analysis of a yeast centromere (CEN3). J. Cell Sci. Suppl. **1**:43–58.

12. Carthew, R. W., L. A. Chodosh, and P. A. Sharp. 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43**:439–448.

13. Chodosh, L. A., S. Buratowski, and P. A. Sharp. 1989. A yeast protein possesses the DNA-binding properties of the adenovirus major late transcription factor. Mol. Cell. Biol. **9**:820–822.

14. Clarke, L., and J. Carbon. 1983. Genomic substitutions of centromeres in *Saccharomyces cerevisiae*. Nature (London) **305**:23–28.

15. Cumberledge, S., and J. Carbon. 1987. Mutational analysis of meiotic and mitotic centromere function in *Saccharomyces cerevisiae*. Genetics **117**:203–212.

16. Diffley, J. F. X., and B. Stillman. 1989. Similarity between the transcriptional silencer binding proteins ABF1 and RAP1. Science **246**:1034–1038.

17. Dorsman, J. C., W. C. van Heeswijk, and L. A. Grivell. 1988. Identification of two factors which bind to the upstream sequences of a number of nuclear genes coding for mitochondrial proteins and to genetic elements important for cell division in yeast. Nucleic Acids Res. **16**:7287–7301.

18. Drake, J. 1970. The molecular basis of mutation. Holden Day, San Francisco.

19. Gaudet, A., and M. Fitzgerald-Hayes. 1987. Alterations in the A+T-rich region of *CEN3* affect centromere function in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7**:68–75.

20. Gaudet, A., and M. Fitzgerald-Hayes. 1989. Mutations in *CEN3* cause aberrant chromosome segregation during meiosis in *Saccharomyces cerevisiae*. Genetics **121**:477–489.

21. Hartwell, L. H. 1980. Mutants of *Saccharomyces cerevisiae* unresponsive to cell division control by polypeptide mating hormone. J. Cell Biol. **85**:811–822.

22. Hegemann, J. H., R. D. Pridmore, R. Schneider, and P. Philippsen. 1986. Mutations in the right boundary of *Saccharomyces cerevisiae* centromere 6 lead to nonfunctional or partially functional centromeres. Mol. Gen. Genet. **205**:305–311.

23. Hegemann, J. H., J. H. Shero, G. Cottarel, P. Philippsen, and P. Hieter. 1988. Mutational analysis of centromere DNA from chromosome VI of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8**:2523–2535.

24. Hieter, P., D. Pridmore, J. H. Hegemann, M. Thomas, R. W. Davis, and P. Philippsen. 1985. Functional selection and analysis of yeast centromeric DNA. Cell **42**:913–921.

25. Huynh, T. V., R. A. Young, and R. W. Davis. 1984. Constructing and screening cDNA libraries in λgt10 and λgt11, p. 49–78. *In* D. Glover (ed.), DNA cloning techniques: a practical approach, vol. I. IRL Press, Oxford.

26. Ito, H., Y. Fukuda, K. Murata, and A. Kimura. 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **153**:163–168.

27. Jiang, W., and P. Philippsen. 1989. Purification of a protein binding to the *CDEI* subregion of *Saccharomyces cerevisiae* centromere DNA. Mol. Cell. Biol. **9**:5585–5593.

28. Jones, E. W., and G. R. Fink. 1982. Regulation of amino acid and nucleotide biosynthesis in yeast, p. 181–299. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), The molecular biology of the yeast *Saccharomyces*: metabolism and gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

29. Koshland, D., J. C. Kent, and L. H. Hartwell. 1985. Genetic analysis of the mitotic transmission of minichromosomes. Cell **40**:393–403.

30. Kurjan, J., and B. D. Hall. 1982. Mutations at the *Saccharomyces cerevisiae SUP4* tRNA^{Tyr} locus: isolation, genetic fine-structure mapping, and correlation with physical structure. Mol. Cell. Biol. **2**:1501–1513.

31. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.

32. Lewis, A. E. 1966. Biostatistics. Reinhold, New York.

33. Lipman, D. J., and W. R. Pearson. 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.

34. Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262**:10035–10038.

35. McGrew, J., B. Diehl, and M. Fitzgerald-Hayes. 1986. Single-base-pair mutations in centromere element III cause aberrant chromosome segregation in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **6**:530–538.

36. Mortimer, R. K., and D. C. Hawthorne. 1969. Yeast genetics, p. 385–460. *In* A. H. Rose and J. S. Harrison (ed.), The yeasts. Academic Press, Inc., New York.

37. Ng, R., and J. Carbon. 1987. Mutational and in vitro protein-binding studies on centromere DNA from *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7**:4522–4534.

38. Panzeri, L., A. Landonio, A. Stolz, and P. Philippsen. 1985. Role of conserved sequence elements in yeast centromere

DNA. EMBO J. **4:**1867–1874.

39. **Reid, B. J., and L. H. Hartwell.** 1977. Regulation of mating in the cell cycle of *Saccharomyces cerevisiae.* J. Cell Biol. **75:** 355–365.

40. **Rose, M., P. Grisafi, and D. Botstein.** 1984. Structure and function of the yeast *URA3* gene: expression in *Escherichia coli.* Gene **29:**113–124.

41. **Rothstein, R. J.** 1983. One-step gene disruption in yeast. Methods Enzymol. **101:**202–211.

42. **Sawadogo, M., and R. G. Roeder.** 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell **43:**165–175.

43. **Sherman, F., G. R. Fink, and J. B. Hicks.** 1986. Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

44. **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.

45. **Thomas, D., H. Cherest, and Y. Surdin-Kerjan.** 1989. Elements involved in *S*-adenosylmethionine-mediated regulation of the *Saccharomyces cerevisiae MET25* gene. Mol. Cell. Biol. **9:** 3292–3298.

46. **Vinson, C. R., K. L. LaMarco, P. F. Johnson, W. H. Landschultz, and S. L. McKnight.** 1988. *In situ* detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Dev. **2:**801–806.