Mutational analysis of the *Saccharomyces cerevisiae* general regulatory factor CP1

Daniel C. Masison, Kevin F. O'Connell and Richard E. Baker*

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

Received March 26, 1993; Revised and Accepted July 15, 1993

ABSTRACT

The *Saccharomyces cerevisiae* general regulatory factor CP1, a helix-loop-helix protein that binds the centromere DNA element I (CDEI) of yeast centromeres, is required in yeast for optimal centromere function and for methionine prototrophy. Mutant alleles of *CEP1*, the gene encoding CP1, were generated by linker insertion, 5'- and 3'-deletion, and random mutagenesis and assayed for DNA binding activity and their ability to confer CP1 function when expressed in yeast. A heterologous CDEI-binding protein, TFEB, was also tested for CP1 function. The results suggested that DNA binding is required for both biological functions of CP1 but is not sufficient. A direct and quantitative correlation was observed between the chromosome loss and nutritional (*i.e.*, Met) phenotypes of strains carrying loss of function alleles, but qualitatively the chromosome loss phenotype was more sensitive to decreased CP1 expression. The data are consistent with a model in which CP1 performs the same general chromatin-related function at centromeres and *MET* gene promoters and is normally present in functional excess.

INTRODUCTION

The *Saccharomyces cerevisiae* centromere-binding protein CP1, encoded by the gene *CEP1* (also designated *CBF1* and *CPF1*), is a member of the basic region-helix-loop-helix (b-HLH) family of DNA binding proteins and is required in yeast for optimal chromosome segregation and methionine prototrophy (1, 2, 3). Its role in chromosome segregation is mediated through an interaction with the conserved DNA element I (CDEI) found in *S. cerevisiae* centromeres. This degenerate element, RTCACRTG (R=purine), contains the consensus CANNTG motif recognized by all known b-HLH DNA binding proteins. Strains harboring *cep1* null mutations show increases in the frequency of mitotic chromosome nondisjunction and in premature segregation of sister chromatids during meiosis (4). The requirement for CP1 in methionine metabolism appears to be exerted at the level of transcription. Consensus CDEI sites are found in the promoter regions of most methionine biosynthetic (*MET*) genes, and it has

been shown that *cep1* mutants fail to properly express *MET16*, the gene encoding phosphoadenylylsulfate (PAPS) reductase (5).

CP1 does not appear to be a typical transactivator protein (*e.g.*, GAL4, GCN4, PHO4). Thomas *et al.* (5) found that LexA-CP1 fusion proteins fail to activate LexA DNA binding site-driven reporter genes, suggesting that CP1 lacks inherent transcriptional activation function. These authors proposed that CP1 binding at *MET* gene promoters is required to allow binding of MET4, a B-ZIP factor, which appears to be the direct transactivator of *MET* gene transcription. The actual mechanism by which CP1 might carry out this facilitator function is not known, but Mellor *et al.* (3) have shown that binding by CP1 alters the chromatin structure of *MET25*, *GAL2*, and *TRP1* gene promoter regions, all of which contain CDEI sites. Specifically, these regions are less sensitive to digestion by micrococcal nuclease in strains that lack CP1. One implication of these results is that CP1 mediates increased accessibility of *MET* gene promoter regions to interactions by other transcription factors (MET4?), presumably by reorganizing the local chromatin structure (5). At centromeres, CP1 may act similarly to facilitate the binding of other kinetochore proteins such as the CDEIII-associated motor protein CBF3 (6, 7).

CP1 has been grouped with RAP1, REB1, and ABF1, proteins which comprise a family of *S. cerevisiae* factors known as general regulatory factors (GRF's) (8, 9, 10, 11, 12). All of these factors are relatively abundant, sequence-specific DNA binding proteins that have many genomic sites of interaction and are involved in diverse chromosomal functions. In addition to having both positive and negative effects on gene transcription, GRF's have been shown to be associated with chromosomal origins of replication (ABF1), telomeres (RAP1), and centromeres (CP1). REB1 (also known as factor Y or GRF2) is a strong positioner of nucleosomes (13).

Considerable evidence suggests that specific DNA binding is required for CP1 function. The equivalent and nonadditive effect of cis and trans CDEI/*cep1* mutations on mitotic centromere function indicates that CP1 interacts at centromeric CDEI sites (1) in carrying out its kinetochore-related function. Domain swap experiments strongly suggest that direct DNA binding is also essential for the transcription function of CP1. Dang *et al.* (14) have engineered a hybrid CP1 protein in which the basic region

* To whom correspondence should be addressed

of the CP1 b-HLH domain is replaced by that of the mammalian b-HLH protein AP4. The resulting mutant, CP1-AP4 has the DNA binding specificity of AP4 (CAGCTG), is unable to bind CDEI sites (CACGTG), and does not confer methionine prototrophy when expressed in yeast. When Met⁺ revertants of CP1-AP4 are selected, all contain an amino acid change in the basic region (amino acid 235) that also reverts the binding specificity to CACGTG. A seemingly contradictory result has been obtained by Mellor *et al.* who have isolated a CP1 mutant which apparently fails to bind DNA yet confers methionine prototrophy (15). They have concluded that direct DNA binding is required only for the centromere function of CP1 and not the transcription function.

The present study was undertaken as an attempt to resolve the discrepant findings concerning CP1 DNA binding and biological function. *In vitro*-constructed *cep1* mutants were generated, introduced into *cep1* null strains of yeast, and tested for function. The DNA binding affinities and expression levels of the mutant gene products were determined in order to establish quantitative correlations between binding and function. A second objective of the study was to determine if specific DNA binding alone is sufficient for CP1 function. For this, we analyzed the ability of a heterologous CDEI-binding protein to rescue CP1 function in yeast. The results were most consistent with the view that direct DNA binding is essential but not sufficient for both kinetochore and transcription functions of CP1 and that CP1 is normally present in functional excess.

MATERIALS AND METHODS

Strains and media

Strains SMAF13 α [*MAT α* cry1 his4-580 lys2 trp1 SUP4-3 ade2-1 leu2 ura3-52 ade3 *cep1::URA3-11 sma1-F13* (pho80)] and R31-1A [*MAT α* leu2 Δ 1 lys2-801 trp1 Δ 1 ura3-52 ade2-101 his3 Δ 200 CFVII(RAD2.d.URA3 SUP11)] have been described (16). Strain R31-3B is identical to R31-1A except that it also carries the *cep1::TRP1* gene disruption allele (16). All media were as described (4). Yeast transformations were performed using the lithium acetate procedure as modified by Schiestl (17).

CEP1 expression vectors and mutant alleles

Figure 1 shows the two basic vectors used for expressing CP1. Plasmid pDR28ARS contains, in addition to ARS1, CEN3, and the selectable marker *LEU2*, an epitope-tagged *CEP1* gene (*CEP1*^{FLAG}). *CEP1*^{FLAG} differs from wild-type *CEP1* in that codons 1–8 were changed to encode MDYKDDDD which specifies the FLAGTM epitope developed by Immunex Corporation (Seattle) (18). The HindIII site encoded at codons 9/10 of wild-type *CEP1* is retained in the same reading frame in *CEP1*^{FLAG}. The *E. coli* T7 expression vector pJH7 was derived from pET-3d (19) by destroying the XbaI and HindIII sites, ligating an XbaI nonsense linker (encoding translation stops in all three reading frames) to the BamHI site, and inserting *CEP1*^{FLAG}. Yeast multicopy (YEp) constructs were obtained by subcloning *CEP1* BamHI–BglII fragments (from pDR28ARS and derivatives) into the BamHI site of the yeast 2 μ -*LEU2* vector YEp351 (20); all plasmids have the *CEP1* segments in the same relative orientation.

The LI-301 allele was constructed by limited PstI digestion of pDR28ARS, removal of the 3' overhanging ends using Klenow enzyme in the absence of dNTP's, and religation in the presence of 10-bp synthetic XhoI linkers (5'-CCCTCGAGGG-3'). Alleles

Figure 1. Plasmids used for expression of *CEP1* alleles in yeast (pDR28ARS) and *E. coli* (pJH7).

LI-381PG and LI-381SR were constructed by digesting pDR28ARS with MluI, filling in the staggered ends with Klenow polymerase, and ligating in the presence of either 8 bp SmaI linkers (5'-GCCCCGGGC-3') or 8 bp XhoI linkers (5'-CCTCG-AGG-3'), respectively. The LI-259 allele was derived from the primary *CEP1* clone pDR4-2 (1) by partial SspI digestion and ligation in the presence of 12-bp SacII linkers (5'-TCCCCGC-GGGA-3'). One of the ligation products contained a single linker inserted at codon 259. The HindIII–BglII fragment of this plasmid was then used to replace that of pDR28ARS. DNA sequencing was performed on all of the mutant constructs to verify that the linkers had inserted as expected and the *CEP1* reading frame was preserved.

The R235H point mutant was obtained by hydroxylamine mutagenesis of pDR28ARS *in vitro* and screening for a Met[–] phenotype by plasmid shuffling (21, 22). One mutant plasmid contained two point mutations, a G to A transition at nucleotide position 959, and a G to T transversion at position 1175 (nucleotide coordinates of reference (1)). The mutations were separated by subcloning HindIII–PstI and PstI–BglII fragments of pHA4 into pDR28ARS. The mutation at position 1175 was silent, but the mutation at position 959 (R235H) resulted in loss of CP1 function.

The 5' deletions were made by exonuclease III digestion of HindIII-cut pDR4-2 using Erase-a-Base kit components (Promega) following the manufacturers instructions. After addition of HindIII linkers (GAAGCTTC), the DNA was cleaved with BglII and fragments ligated to HindIII–BglII digested pDR28ARS. Deletion endpoints were determined by DNA sequencing. The 3' deletion allele, Δ 302, was obtained in a

similar fashion except exonuclease digestion was initiated at the BglII site, 14-bp XbaI nonsense linkers (having stop codons in all reading frames) were added, and the resulting HindIII–XbaI fragment was initially ligated to HindIII–XbaI-cut pJH7. Subsequently, the CΔ302 segment was removed as a HindIII–BamHI fragment and ligated to HindIII–BglII cut pDR28ARS for expression in yeast. This plasmid transformed yeast poorly, possibly because the *CEP1* transcription unit directly abutted the ARS1 sequence. The problem was alleviated when 0.6 kbp of *CEP1* downstream sequence (along with 19 C-terminal codons of CP1) were reintroduced. The resultant construct, pJH11, is effectively an internal deletion of *CEP1* codons 303–332, but translation terminates at codon 302 due to the XbaI nonsense linker at the deletion endpoint.

TFEB expression plasmids all contain nucleotides 871 (RsaI) to 1523 (SmaI) of the TFEB cDNA isolated by Carr and Sharp (23). Using standard procedures, an 8-bp HindIII linker was added at the 5'-end and a 14-bp XbaI nonsense linker added at the 3'-end. This segment was then inserted into pJH7 for *E. coli* expression and into the pJH11 vector backbone for expression in yeast.

Assays of *cep1* phenotypes

Mitotic chromosome loss rates were measured by fluctuation analysis as described (24) except that leucine was omitted from all media to maintain selection for the *CEP1*-containing plasmids. The tester strain (R31-3B) carried the nonessential chromosome fragment *CFVII(RAD2.d.URA3 SUP11)* whose rate of loss was measured. A semi-quantitative assay was devised to score methionine-independent growth. Several independent transformant colonies of each strain were grown as patches of cells at 30°C on plates lacking leucine, then replica plated in triplicate onto ⁻Leu dropout and ⁻Leu ⁻Met double dropout plates. The plates were incubated at 22°C, 30°C, and 34°C and scored after 24, 48, and 72 hours. On each day, the patches were assigned a score of 2 (confluent growth), 1 (less than confluent growth) or 0 (no detectable growth), and the individual scores added after 3 days. Thus, the maximum score is 6 and the minimum 0. Wild-type cells routinely give a score of 6 and *cep1* mutants a score of 0.

Preparation of yeast and *E. coli* extracts

Yeast strains (R31-3B) carrying CP1 expression plasmids were grown at 30°C in 100 ml of selective (⁻Leu) medium to mid log phase. All subsequent manipulations were carried out on ice or in a 4°C cold room. The cells were washed once with water, resuspended in two pellet volumes of 200 mM Tris (pH8.1)-10 mM MgCl₂-10 mM β-mercaptoethanol-10% glycerol-1 mM PMSF, transferred to 2-ml screw cap microfuge tubes containing 1 ml of 0.5-mm glass beads, and broken by two-20 s bursts in a Biospec Products Mini-Bead Beater. Tubes were cooled for 20 s in an ice-ethanol bath between agitations. Lysates were transferred to 1.6-ml tubes and spun in a microcentrifuge for 10 min to remove cellular debris. Supernatants were frozen in aliquots on dry ice and thawed only once before use. Protein concentrations were determined by the method of Bradford using bovine serum albumin as standard (25).

Soluble *E. coli* extracts were prepared essentially as described by Studier et. al. (19). The host strain for T7-induced expression was BL21(DE3) carrying the plasmid pLysS. BL21(DE3) transformants carrying pJH7-derived CP1 expression plasmids were grown overnight at 37°C in LB broth containing ampicillin

and chloramphenicol. The overnight cultures were diluted 1:100 into fresh medium and grown at 37°C to an OD₆₀₀ of 0.3–0.6 at which time isopropylthiogalactoside (IPTG) was added to a final concentration of 0.4 mM. Incubation was continued for an additional 90 min. Cells were pelleted, resuspended in extract buffer [200 mM Tris (pH8.1), 10 mM MgCl₂, 10 mM β-mercaptoethanol, 10% glycerol, 0.5 mM PMSF, 0.5 mM DTT, 0.2% NP40] to a final concentration of 4 OD₆₀₀ units per ml, and chilled on ice. Cell breakage was accomplished by sonication using a Branson J-17A sonicator at full power for two 30 s bursts with a 30 s rest on ice between disruptions. After pelleting cellular debris by centrifugation for 10 min at 4°C in a microcentrifuge, the supernatants were transferred to clean tubes and frozen on dry ice.

DNA binding assays

Probe for DNA binding assays was prepared from plasmid pDR1-7 which contains a single copy of the CDEI oligonucleotide 5'-GATCCAAATAAGTCACATGATGATA(GATC)-3' cloned into the BamHI site of pUC18. After EcoRI/HindIII digestion and Klenow fill-in using [α-³²P]dATP, the 72-bp probe was isolated by polyacrylamide gel electrophoresis and electro-elution. Probe DNA concentration was determined using a Hoeffer Scientific model TKO 100 DNA fluorimeter. Alternatively or in addition, samples of probe DNA were run on a 6% polyacrylamide gel along with known amounts of EcoRI–HindIII digested pDR1-7 plasmid DNA and DNA concentration determined after ethidium bromide staining (26).

Gel shift DNA binding assays were performed as described (27). Binding reactions contained (in 15 μl) 20 mM Hepes, 150 mM KCl, 5 mM MgCl₂, 2.5% glycerol, 0.25 mg/ml bovine serum albumin, 0.05% Nonidet P-40, 20,000 cpm of ³²P-labeled probe DNA (200–500 pg), unlabeled competitor DNA, and either yeast or *E. coli* extract. *E. coli* extract assays used 1 μl of a 1:100-diluted extract (30–60 ng total protein) and 2 μg of poly d(I-C) as unlabeled competitor. Yeast extract assays used 20 μg total yeast protein (2–5 μl extract) and either 2.5 μg pDR1-12 plasmid DNA (specific competitor) or 2.5 μg pDR2-14 plasmid DNA (nonspecific competitor). The specific competitor DNA was plasmid pDR1-12 which contains a pentamer of the CDEI oligo described above in the BamHI site of pUC18. The amount used (2.5 μg/reaction) provided a 500- to 1300-fold molar excess of CDEI sites relative to the labeled probe. Nonspecific competitor was plasmid pDR2-14 which is identical to pDR1-12 except that it contains a pentamer of an unrelated oligonucleotide.

Determination of relative apparent binding affinities

Relative apparent binding affinities were determined by quantitative gel shift DNA binding assays of *E. coli*-produced gene products. To quantitate the amount of binding protein in each extract, equal volumes of extract were run on 12% SDS polyacrylamide (Laemmli) gels and the gels stained with Coomassie blue. The CP1 bands were clearly detectable above the background of *E. coli* proteins. After destaining, the gels were sandwiched between two sheets of 3M transparency film and scanned using a Hoeffer Scientific GS-300 Densitometer. Peak integration was accomplished using GS-365 PS Electrophoresis Data System software (Hoeffer). Gel lanes of extracts of *E. coli* not carrying CP1 expression plasmids were scanned to determine background staining in the regions of the gels containing CP1 products. This background accounted for less than 10% of the integrated CP1 signal in all cases except for LI-318SR which

was poorly expressed. Background-subtracted CP1 integration values were corrected for differing polypeptide molecular weight and then normalized to the CP1 signal obtained from a single extract of the NΔ208 mutant, an aliquot of which was included on all gels for standardization. Four separate scans of samples from at least two independent extracts were carried out for each CP1 protein. Results from duplicate scans differed by less than 10%.

The absolute concentration of CP1 in the extracts was estimated by running the 'standard' NΔ208 extract on a gel along with known amounts of homogeneously pure human serum albumin (a gift of R. Feldhoff, University of Louisville School of Medicine). After densitometric analysis as described above, the concentration of NΔ208 polypeptide was determined from a standard curve. The 'standard' extract contained 0.47 mg/ml of NΔ208 polypeptide (29 μM).

After performing gel shift DNA binding assays as described above, dried gels were analyzed on a Betagen Betascope 603 blot analyzer to quantitate the amount of bound and free probe in each binding reaction. Relative apparent CP1 binding constants (K_{rel}) were calculated using the following equation:

$$K_{rel} = \frac{K_{test}}{K_{ref}} = \frac{R_{test}}{R_{ref}} \cdot \frac{[CP1]_{ref}}{[CP1]_{test}} \quad (1)$$

K_{test} and K_{ref} are the apparent binding constants of, respectively, the test and reference proteins; R_{test} and R_{ref} are the ratios of bound to free probe for the respective binding reactions; $[CP1]_{test}$ and $[CP1]_{ref}$ are the total concentration of test and reference proteins in the reactions, respectively. The standard NΔ208 extract was used as the 'reference' in all cases. Equation 1 assumes that the concentration of unbound CP1 protein in each binding reaction is approximately equal to the input concentration of CP1, *i.e.*, conditions of protein excess. This condition was met in our reactions where the total CP1 concentration was 10–20 nM and the concentration of bound probe was 0.01–0.15 nM. The reported results (Table 1) are the means of determinations for two or three independent extracts of each protein. The average deviation from the mean was less than 20%.

Immunodetection of FLAGged CP1 proteins

Crude yeast proteins (up to 30 μg of total protein) were resolved on 12% polyacrylamide SDS (Laemmli) minigels (0.75 mm). Proteins were transferred to Immobilon PVDF membranes (Millipore) in 1×CAPS buffer containing 10% methanol (28). Detection of FLAGged proteins was carried out using the Amersham ECL detection system following the manufacturers guidelines. The primary antibody used was monoclonal anti-FLAG-M5 mouse IgG1 diluted to 10 mg/ml in 0.1%TTBS [0.1%Tween-20, 20 mM Tris–HCl (pH7.5), 0.5 M NaCl]. The secondary antibody was peroxidase-conjugated goat anti-mouse Ig (Boehringer Mannheim Biochemicals) diluted 1:3,000 in 0.1%TTBS containing 2.5% dry milk. After reaction with antibody, membranes were washed three times in 0.1%TTBS-0.1% dry milk.

RESULTS

cep1 mutations and phenotype analysis

The *cep1* mutants chosen for the first phase of this study are diagrammed schematically in Figure 2. Mutations CA302, LI-301, LI-318PG, and LI-318SR disrupt the C-terminal region of CP1. It was known from preliminary experiments that these mutations reduced but did not abolish specific DNA binding and might therefore be useful for correlating binding and function. For comparison, two non-DNA binding mutants were also analyzed: mutation LI-259 is a four codon insertion into helix 2 of the b-HLH domain; mutation R235H is a point mutation in the basic region. The last allele shown in Figure 2 was derived from a cDNA clone of the mammalian b-HLH protein TFEB (23). This factor, which has a binding specificity very similar to CP1 (23), was tested to determine if a heterologous CDEI-binding protein would confer CP1 function in yeast. For positive controls we used both wild-type *CEP1* as well as the N-terminal truncation mutant NΔ208 (Figure 6), which was known to have wild-type function (3). All of the alleles encoded an epitope tag (FLAG) at the protein N-terminus but retained native *CEP1* promoter and translation initiation sequences. Substituting the eight amino acid FLAG for the natural CP1 N-terminus leads to no measurable alteration in CP1 function (data not shown).

Table 1. Phenotypes conferred by *cep1* mutant alleles

Allele	Met Phenotype YCp	YEp	CFVII Loss Rate ^a Absolute ^b	Relative	CDEI Binding ^c
WT	6	6	3.0×10^{-4}	1.0	1.0
NΔ333 ^d (30°)	0	0	1.2×10^{-2}	40	
NΔ333 (22°)	0	0	1.2×10^{-2}	37	<0.002
NΔ208	6	6	1.4×10^{-4}	0.5	0.5
CA302	0	0	1.1×10^{-2}	37	0.016
LI-301	0	0	9.3×10^{-3}	31	0.024
LI-318PG (30°)	1	6	5.0×10^{-3}	17	
LI-318PG (22°)	3	6	6.2×10^{-3}	21	0.13
LI-318SR	6	6	3.2×10^{-4}	1.1	0.14
LI-259	0	0	1.0×10^{-2}	33	<0.002
R235H	0	0	9.3×10^{-3}	31	<0.002
TFEB	0	0	9.0×10^{-3}	30	0.5

^aDetermined by fluctuation analysis using YCp transformants.

^bLoss events/cell/division.

^cRelative apparent binding affinities determined using *E. coli*-produced proteins (see Materials and Methods) and expressed as normalized values.

^dNΔ333 lacks all but 17 C-terminal CP1 codons (Figure 6) and is used as the negative control.

Before testing function in yeast, the mutant *cep1* alleles were expressed in *E. coli* to obtain accurate estimates of the relative apparent CDEI-binding affinities of the gene products. The apparent affinities varied over a wide range (Table 1). The NΔ208 and TFEB gene products bound with an apparent affinity only 2-fold lower than that of wild-type CP1. In contrast, the R235H and LI-259 gene products, both altered in the CP1 b-HLH domain, possessed no detectable CDEI-binding activity. The C-terminal domain mutants had significantly reduced but measurable DNA binding activities: the LI-318 mutations resulted in about a 9-fold reduction in apparent binding affinity; while the LI-301 and CΔ302 mutations reduced apparent binding affinity 40- and 60-fold, respectively.

Next, the alleles were introduced into *cep1* null yeast strains on both single- and multicopy plasmids and tested for expression. An indirect test of expression was made by measuring CDEI-binding activity present in crude extracts (Figure 3). Equal amounts of crude yeast protein were used for the assays; therefore, the amount of specifically bound probe depends both upon binding affinity and upon the relative concentration of binding protein in the extract. For the most part, the activities

Figure 2. *CEPI* mutant alleles. Alleles were constructed as described in Materials and Methods. Numbers below the diagrams indicate codon positions. The 3' deletion, CΔ302, is actually an internal deletion (codons 302–332) terminated by an XbaI nonsense linker (cross hatched box) after codon 301. LI-301 is a linker insertion at the PstI site which replaces amino acids 302 and 303 (LQ) with those indicated. The remaining linker insertions contain in-frame insertions of four codons immediately 3' to the codon positions indicated. R235H is a histidine to arginine substitution at position 235 (basic region). The TFEB allele expresses the DNA binding domain of the mammalian factor TFEB (23). The approximate positions of the heterologous b-HLH domain and the leucine zipper (ZIP) of TFEB are indicated.

Figure 3. DNA binding activity of yeast transformants carrying mutant *cep1* alleles. Gel shift DNA binding assays of crude extracts were carried out as described in Materials and Methods. The addition of specific (+) or nonspecific (–) unlabeled competitor is indicated. The various mutations are described in Figures 2 and 6. The YEp351 transformant carries the expression vector (*LEU2*) without a CP1 insert. The assay shown utilized extracts prepared from *cep1* transformants carrying the mutant alleles on multicopy (YEp351-derived) plasmids. The same relative results were obtained when extracts were prepared from transformants carrying single-copy plasmids, but the absolute levels of CDEI-binding activity were less (not shown).

Figure 4. Immunodetection of epitope-tagged proteins. Immunoblotting (see Materials and Methods) was carried out using the same crude yeast extracts that were used for the binding assay shown in Figure 3. Thirty μ g of total protein were loaded except for extracts of strains carrying LI-259 and R235H where only 5 μ g were loaded and strain R19-4 where 80 μ g was used. R19-4 carries a chromosomal copy of *CEPI* ^{FLAG}. The two left-hand lanes contain approximately 1 ng of *E. coli*-produced CP1 (pJH7) loaded alone (lane 2) or mixed with control (YEp351) yeast extract (lane 1). The gel was intentionally overloaded and the blot overexposed to maximize sensitivity. (Notice the absence of background bands in the R235H and LI-259 lanes where less total protein was loaded.)

observed corresponded to those that would be expected from the measured relative binding affinities of the respective gene products: wild-type activity was observed for NA208 and TFEB; reduced binding was observed for LI-318SR and LI-318PG; and no binding was detected for R235H and LI-259. In all cases where binding was observed, it was specific, as binding to the probe was reduced in the presence of excess unlabeled CDEI site-containing DNA. The binding activities observed for LI-301 and CA302 were not in line with expectations. Given the 50-fold-reduced binding affinity of these mutant gene products, LI-301 and CA302 transformants were expected to yield extracts with extremely low (or undetectable activity). No detectable binding was observed for CA302, but nearly wild-type activity was observed for LI-301.

Immunoblot analysis of the extracts (Figure 4) provided a direct estimate of mutant gene expression and resolved the discrepancy with respect to LI-301 and CA302. The LI-301 gene product, as well as those of R235H and LI-259, were expressed at greater than wild-type levels. The high relative concentration of LI-301 protein probably explains why appreciable CDEI-binding activity was detected in LI-301 extracts despite the reduced binding affinity. A gene product was not detected for CA302. Apparently CA302 protein is not stable in yeast (although it is stable in *E. coli*), making this mutant uninformative. Interestingly, the LI-318PG product was also not detected, even though CDEI-binding activity was present in the extract. Evidently, the DNA binding assays were more sensitive at detecting CP1 than immunoblotting, and the concentration of CP1 product in the LI-318PG extract lay between the detection limits of the two assays. In any event, this result suggested that the steady-state level of LI-318PG gene product was less than that of LI-318SR and that the apparent difference in DNA binding activity in these two extracts (Figure 3) was real. The NA208, TFEB, and LI-318SR gene products were detected at levels comparable to wild-type CP1.

The results of CP1 functional assays are reported in Table 1. With two exceptions, CP1 function with respect to both methionine biosynthesis and mitotic chromosome segregation correlated with the presence or absence of CDEI-binding activity. The nonbinding mutants R235H and LI-259 lacked any measurable CP1 function. The NA208 and LI-318SR gene products conferred wild-type function. The LI-318PG allele did not confer full function when present on a single-copy (YCp) plasmid; however, the Met phenotype was partially corrected at reduced temperature (22°C), and both phenotypes were rescued to wild-type (Met) or near wild-type (CF loss) levels when LI-318PG was expressed from a multicopy (YEp) plasmid. [Mitotic CF loss rate was rescued to 8.3×10^{-4} events/division.] The temperature sensitivity and dosage dependence of the LI-318PG phenotype suggested that the primary defect was in expression, possibly due to protein instability, rather than a specific deficiency in function. The reduced level of binding activity and immunoreactive protein relative to LI-318SR (Figures 3 and 4) corroborate this conclusion. Despite having wild-type CP1 function, multicopy plasmid transformants of LI-318SR and LI-318PG expressed much lower than wild-type levels of CDEI-binding activity (Figure 3), suggesting that CP1 is normally present in significant excess.

The exceptions to the correlation between CDEI-binding activity and CP1 function were TFEB and LI-301. The TFEB gene product, which was expressed at levels similar to and had an apparent CDEI-binding affinity only 2-fold lower than wild-

Figure 5. Dominant negative effects of *cep1* mutant alleles. Transformants carrying the indicated alleles on single- (YCp) or multicopy (YEp) plasmids were grown to mid log phase in medium lacking leucine (to select for the expression plasmid), centrifuged, and diluted to a final concentration of 2×10^6 cells per ml in medium lacking both leucine and methionine. For each strain, a series of five, 5-fold serial dilutions were prepared in the double selection medium and 5 μ l aliquots were spotted onto single selection (-Leu) and double selection (-Leu -Met) plates. Plates were incubated at 30°C and photographed after 48 (*CEP1*) or 72 hours (*cep1 sma1*). The YE p351 transformant carries the expression vector (*LEU2*) lacking a CP1 insert.

type CP1, exhibited no detectable *cep1*-complementing activity under any condition tested. Likewise, LI-301 was totally inactive in rescuing CP1 function, even though multicopy expression resulted in almost wild-type levels of CDEI-binding activity in crude extracts.

Dominant negative effects

The finding that TFEB and LI-301 transformants expressed wild-type levels of CDEI-binding activity but were devoid of CP1 function strongly implied that CDEI-binding activity *per se* is not sufficient to confer CP1 function. But do TFEB and LI-301 gene products actually interact with CDEI sites *in vivo*? If so, these mutants should display dominant negative phenotypes, *i.e.*, they should inhibit the function of wild-type CP1 by nonproductively occupying CDEI sites. When tested, none of the mutants conferred a trans dominant phenotype in a *CEP1*⁺ host; however, TFEB and LI-301 did exhibit dominant negative Met phenotypes when expressed in a *cep1 pha80* strain background.

Mutations in *pha80* suppress *cep1* methionine auxotrophy (16). The suppression appears to be due to a direct activation of methionine biosynthetic genes by the positive transcription factor PHO4, which shares binding specificity with CP1 (14, 29). The PHO4 activation pathway is not available under normal conditions, because PHO4 is kept inactive by the negative regulator PHO80 (30, 31). Mutations in *pha80* free PHO4 to act constitutively, and *cep1 pha80* double mutants are phenotypically Met⁺ (16). The test for dominant negative Met phenotypes in the *cep1 pha80* (*sma1*) background is shown in Figure 5 (lower panels). [The *sma1* (suppressor of *cep1* methionine auxotrophy) complementation group is allelic to *pha80*.] When carried on either single- or multicopy vectors, LI-301 and TFEB caused dramatic reductions in the ability of

		bp to nearest in-frame ATG	Methionine phenotype	CFVII Loss Rate	
				Absolute	Relative
WT	GAT GAT <u>AAG CTT</u> TCT ACT (+1)	–	6	3.0×10^{-4}	1.0
NΔ33	GAT GAT aag ctt cGG CCA GGA (+3)	359	1	1.1×10^{-2}	34
NΔ86	GAT GAT aag ctt cGA A GA TGC (+2)	199	5	9.4×10^{-4}	3
NΔ137	GAT GAT aag ctt cAA CCC CAT (+3)	47	2	8.0×10^{-3}	27
NΔ162	GAT GAT aag ctt cGC AAT CTA (+3)	17	4	1.6×10^{-3}	5
NΔ208	GAT GAT aag ctt cAA CCT ACT (+1)	–	6	1.4×10^{-4}	0.5
NΔ333	GAT GAT aag ctt cAC TCA CAA (+3)	–	0	1.2×10^{-2}	40

Figure 6. *cep1* frame-shift alleles and their ability to complement *cep1* phenotypes. The fusion joint at the deletion endpoint is shown above each diagram. HindIII linker sequences are in lower case, FLAG and *CEP1* sequences in upper case. The naturally occurring HindIII site spanning codons 9/10 in the wild type allele (WT) is underlined. Nucleotides are grouped as triplets to indicate the reading frame established by the ATG initiator codon provided by the expression vector. Underlined triplets indicate the correct CP1 reading frame with the numbers in parentheses indicating the net shift in reading frame relative to the initiating ATG (and FLAG). The alleles are named to indicate the native *CEP1* codon at the fusion site. Methionine phenotype and CFVII loss rates were determined as described in Materials and Methods. Absolute loss rates are loss events/cell/division.

the respective strains to grow in the absence of methionine. The non-DNA binding mutants R235H and LI-259 had no negative effects on the growth of the *cep1 smal* strain in the absence of methionine (neither do they exhibit negative phenotypes in a *CEP1*⁺ background; Figure 5, top). The inhibitory effects of TFEB and LI-301 were not due to a direct inactivation of PHO4, because *PHO5* expression in the *cep1 smal* transformants was unaffected (not shown). [PHO5, encoding repressible acid phosphatase (rAPase), is a downstream target of PHO4 (31).] That expression of TFEB and LI-301 inhibits suppression by PHO4 implies that both TFEB and LI-301 gene products enter the nucleus and occupy CDEI sites.

CEP1 frame-shift mutants

In constructing N-terminal truncation mutants of CP1, a number of constructs were generated which did not maintain translational reading frame, *i.e.*, *CEP1* coding sequences were out of frame relative to the initiator ATG codon of the expression vector (see Materials and Methods). To our surprise, several such frame-shifted constructs were found to confer a partial Met⁺ phenotype. Apparently, low levels of truncated but functional CP1 gene product were produced either by frame-shift suppression or translational (re)initiation at alternative ATG codons. While the exact mechanism of expression was not known, preliminary characterization of the mutants indicated a wide range of activities, providing us the opportunity to systematically compare Met and chromosome loss phenotypes under conditions of limiting CP1 expression.

The frame-shift alleles and their ability to rescue *cep1* phenotypes are shown in Figure 6. The *CEP1* open reading frame contains in-frame methionine codons at amino acid positions 153

Figure 7. CDEI binding activity of strains expressing *cep1* frame-shift alleles. Binding assays were carried out as described in Materials and Methods. The crude extracts were prepared from RB31-3B transformants carrying the mutant alleles on single-copy (YCp) plasmids. Specific (+) or non-specific (–) plasmid competitor DNA (2.5 μg) was included in each reaction.

and 168. With the exception of NΔ86, the shorter the distance between the deletion endpoint and the nearest in-frame ATG (codon 153 for NΔ33, NΔ86, and NΔ137; codon 168 for NΔ162), the better the allele at complementing *cep1* defects. NΔ86 differs in that it is the best at complementing *cep1* defects and that its CP1 coding sequences lie in the +2 reading frame relative to the initiator ATG and FLAG codons (the other frame-shift alleles lie in the +3 reading frame). The FLAG codons contain three tandem ATG's in the +2 reading frame, one of which may serve as the initiator for NΔ86. Most importantly,

the results in Figure 6 show a direct and quantitative correlation between the ability of an allele to promote methionine-independent growth and its ability to rescue the *cep1* mitotic chromosome loss defect.

To determine if the correlation between DNA binding activity and biological function also held for these frame-shift mutants, extracts of transformants carrying the frame-shift alleles were assayed for CDEI-binding activity. As shown in Figure 7, the only extracts in which specific CDEI binding activity was detected were those of strains carrying the wild-type, NΔ208, and NΔ86 alleles. These three alleles were also the best at complementing *cep1* phenotypes. The NΔ208 extract had wild-type binding activity, while an equal amount of NΔ86 extract protein bound markedly less probe, consistent with a reduced expression of this allele due to the frame-shift mutation. The NΔ86 complex was of a size consistent with it being initiated from an ATG very near the fusion site (*i.e.*, within FLAG). No CDEI-specific binding activity was detected in extracts prepared from transformants carrying the other out-of-frame alleles, consistent with their reduced abilities to rescue *cep1* phenotypes. The apparent lack of binding activity in NΔ33, NΔ137, and NΔ162 extracts probably reflects the detection limit of the assay. There is no reason *a priori* to assume that these gene products would lack CDEI-binding activity. In-frame truncation alleles NΔ80, NΔ182, and NΔ192 (as well as NΔ208) all produce wild-type levels of CDEI-binding activity when expressed from single-copy plasmids (data not shown). Evidently, extremely low levels of CP1 expression suffice for at least partial CP1 function.

DISCUSSION

As mentioned in the Introduction, a preponderance of evidence indicates that direct and specific DNA binding is required for CP1 biological function. Our results fully support that view. With the exception of mutants which bind DNA but are nonfunctional (LI-301, TFEB), the ability of mutant alleles to rescue both *cep1* phenotypes correlates with the expression of CDEI-binding activity. Nonbinding mutants (R235H, LI-259) have no detectable function; mutants expressing wild-type (NΔ208) or reduced binding activity (LI-318SR, multicopy LI-318PG) display wild-type function; those expressing very low (NΔ86) or undetectable (NΔ162, NΔ137, and NΔ33) binding activity have impaired function. The lack of observed activity in the case of the latter three frame-shift mutants probably reflects assay insensitivity, since there is no reason *a priori* to predict that the gene products intrinsically lack DNA binding activity. None of the mutations separated the segregation and transcription functions of CP1. While certain mutants restored methionine prototrophy without completely rescuing the chromosome loss defect (Met⁺Chl⁻), no mutations causing the reverse phenotype (Met⁻Chl⁺) were obtained. In all cases, mutations that had any measurable phenotypic effect invariably affected both functions, and for the N-terminal truncation/frame-shift series, a direct and quantitative correlation was observed in the degree to which Chl and Met phenotypes were complemented. Given the limited number of mutations in hand, we are wary of overinterpreting this result, but the data are consistent with CP1 performing a similar function at *MET* gene promoters and centromeres.

Mellor *et al.* (15) have proposed that CP1-CDEI-binding is required for the centromere activity of CP1 but not to support methionine prototrophy, implying that CP1 acts via different mechanisms to carry out its multiple functions. This conclusion

was based on the analysis of a site-directed basic region *cpf1* (*cep1*) mutant, YAG214, which was found to possess no detectable CDEI-binding activity *in vitro*, but which conferred methionine prototrophy and partially rescued centromere function when expressed in yeast. This is the same phenotype we observe with the NΔ162, NΔ137, and NΔ33 frame-shift alleles: strains grow in the absence of methionine, show intermediate chromosome loss rates, and yield crude extracts devoid of detectable CDEI-binding activity. Our interpretation of this result is simply that very low levels of CP1 activity are sufficient for partial function and that less CP1 is required to maintain methionine prototrophy than to maintain full centromere function. We suggest that the Met⁺Chl⁻ phenotype of the YAG214 mutant could be explained similarly, *i.e.*, the YAG214 gene product has a low but nonzero binding affinity and provides residual function *in vivo*.

CP1 is a relatively abundant protein (27, 35), and results here suggest that the nuclear concentration of CP1 is, in fact, well in excess of the minimum required for wild-type function (at least with respect to the two phenotypes we analyze). First, the LI-318SR allele in single copy is wild-type in function even though its apparent CDEI binding affinity is reduced 9-fold. Second, the N-terminal truncation/frame shift alleles, which appear to be expressed at a very small fraction of wild-type levels, still provide significant CP1 function. Finally, the absence of dominant negative effects in a wild-type *CEP1* background could be interpreted to suggest excess CP1. A cellular surplus of CP1 would explain the previous finding that a centromeric CDEI mutation which reduces binding 45-fold causes no apparent phenotypic effect (27, 36).

Previous mutational analyses of CP1 have identified critical functional domains. As for other members of this family, the b-HLH domain is essential for DNA binding (3, 15, 32). Two of the mutations studied here are located in this domain, and both reduce CDEI-binding affinity by at least 500-fold. The R235H point mutation changes a conserved arginine residue in the basic region which has been shown to be responsible for specific recognition of the inner dinucleotide of the consensus CANNTG binding site motif (CACRTG in the case of CP1) (14, 33). The LI-259 mutation disrupts the second helix of the HLH domain and would be expected to affect CP1 dimerization (32). Dowell *et al.* have reported that the C-terminal region of CP1 (spanning codons 292–320) comprises a second dimerization domain that is essential for DNA binding and CP1 function *in vivo* (34). The significantly decreased DNA binding affinities of the CA302, LI-301, LI-318PG, and LI-318SR proteins confirm the importance of this domain for DNA binding; however, since the CA302 allele lacks two thirds of these amino acids and still binds DNA specifically, we conclude that the HLH domain of CP1 is sufficient to mediate dimerization. In addition to an effect on DNA binding, the LI-318PG mutation appears to destabilize CP1 *in vivo*. This mutation inserts a proline into a region of CP1 which is predicted to be α -helical in structure (34).

While DNA binding appears to be necessary for CP1 function, CDEI-binding *per se* is not sufficient. The TFEB protein has an apparent CDEI binding affinity close to that of wild-type CP1, is expressed at a comparable level, but is devoid of biological function. Also, the inactivity of LI-301 cannot be explained solely by its reduced affinity, since strains expressing LI-301 from a multicopy vector contain wild-type levels of CDEI-binding activity. Both TFEB and LI-301 gene products apparently enter the nucleus and occupy CDEI sites, because both alleles exert

strong dominant negative effects on the ability of PHO4 to activate *MET* genes. Our working hypothesis is that CP1 acts primarily to configure chromatin into a conformation which facilitates the assembly of functional kinetochores and active transcription complexes. CP1 is known to influence chromatin structure at both gene promoter regions and centromeres (3), but it is not known if the CP1-associated chromatin structures are established independently by CP1 or require the participation of cofactors, which could differ in each case. Since CDEI site occupation alone appears to be insufficient to mediate CP1 function, CP1 must contribute an additional activity. The protein domain conferring the effector function is apparently located within the C-terminal third of the protein and may be the target of the LI-301 mutation. This functional domain might interact with histones or other chromatin proteins to enhance (or inhibit) the assembly of higher order structures.

ACKNOWLEDGEMENTS

The authors thank Cynthia Carr and Phil Sharp for donating the TFEB cDNA clone and Keming Zhang for contributing several plasmid constructs. This work was supported by a grant from the National Institutes of Health (GM38566).

REFERENCES

- Baker, R.E. and Masison, D.C. (1990) *Mol. Cell. Biol.*, **10**, 2458–2467.
- Cai, M. and Davis, R.W. (1990) *Cell*, **61**, 437–446.
- Mellor, J., Jiang, W., Funk, M., Rathjen, J., Barnes, C.A., Hinz, T., Hegemann, J.H. and Philippsen, P. (1990) *EMBO J.*, **9**, 4017–4026.
- Masison, D.C. and Baker, R.E. (1992) *Genetics*, **131**, 43–53.
- Thomas, D., Jacquemin, I. and Surdin-Kerjan, Y. (1992).
- Lechner, J. and Carbon, J. (1991) *Cell*, **64**, 717–725.
- Hyman, A.A., Middleton, K., Centola, M., Mitchison, T.J. and Carbon, J. (1992) *Nature*, **359**, 533–536.
- Buchman, A.R., Kimmerly, W.J., Rine, J. and Kornberg, R.D. (1988) *Mol. Cell. Biol.*, **8**, 210–225.
- Devlin, C., Tice-Baldwin, K., Shore, D. and Arndt, K.T. (1991) *Mol. Cell. Biol.*, **11**, 3642–3651.
- Ju, Q., Morrow, B.E. and Warner, J.R. (1990) *Mol. Cell. Biol.*, **10**, 5226–5234.
- Chasman, D.I., Lue, N.F., Buchman, A.R., LaPointe, J.W., Lorch, Y. and Kornberg, R.D. (1990) *Genes and Dev.*, **4**, 503–514.
- Buchman, A.R. and Kornberg, R.D. (1990) *Mol. Cell. Biol.*, **10**, 887–897.
- Fedor, M.J., Lue, N.F. and Kornberg, R.D. (1988) *J. Mol. Biol.*, **204**, 109–127.
- Dang, C.V., Dolde, C., Gillison, M.L. and Kato, G.J. (1992) *PNAS*, **89**, 599–602.
- Mellor, J., Rathjen, J. and Dowell, S.J. (1991) *NAR*, **19**, 2961–2969.
- O'Connell, K.F. and Baker, R.E. (1992) *Genetics*, **192**, 63–73.
- Schiestl, R.H. and Gietz, R.D. (1989) *Curr. Genet.*, **16**, 339–346.
- Hopp, T.P., Prickett, K.S., Price, V.L., Libby, R.T., March, C.J., Cerretti, D.P., Urdahl, D.L. and Conlon, P.J. (1988) *Biotechnology*, **6**, 1204–1210.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff. (1990) *Meth. Enzymol.*, **185**, 60–88.
- Hill, J.E., Myers, A.M., Koerner, T.J. and Tzagoloff, A. (1986) *Yeast*, **2**, 163–167.
- Schatz, P.J., Solomon, F. and Botstein, D. (1988) *Genetics*, **120**, 681–695.
- Kranz, J.E. and Holm, C. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 6629–6633.
- Carr, C.S. and Sharp, P.A. (1990) *Mol. Cell. Biol.*, **10**, 4384–4388.
- Koshland, D. and Hieter, P. (1987) *Meth. Enz.*, **155**, 351–372.
- Bradford, M.M. (1976) *Anal. Biochem.*, **72**, 248–254.
- Baker, R.E., Gabrielson, O. and Hall, B.D. (1986) *J. Biol. Chem.*, **261**, 5275–5282.
- Baker, R.E., Fitzgerald-Hayes, M. and O'Brien, T.C. (1989) *J. Biol. Chem.*, **264**, 10843–10850.
- Matsudaira, P. (1987) *J. Biol. Chem.*, **262**, 10035–10038.
- Fisher, F. and Goding, C.R. (1992) *EMBO*, **11**, 4103–4109.
- Yoshida, K., Ogawa, N. and Oshima, Y. (1989) *Mol. Gen. Genet.*, **217**, 40–46.
- Vogel, K. and Hinnen, A. (1990) *Mol. Microbiol.*, **4**, 2013–2017.
- Jones, N. (1990) *Cell*, **61**, 9–11.
- Van Antwerp, M.E., Chen, D.G., Chang, C. and Prochownik, E.V. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 9010–9014.
- Dowell, S.J., Tsang, J.S.H. and Mellor, J. (1992) *NAR*, **20**, 4229–4236.
- Bram, R.J. and Kornberg, R.D. (1987) *Mol. Cell. Biol.*, **7**, 403–409.
- Gaudet, A. and Fitzgerald-Hayes, M. (1989) *Genetics*, **121**, 477–489.