MOLECULAR AND CELLULAR BIOLOGY, Jan. 1987, p. 403–409
0270-7306/87/010403-07$02.00/0

Vol. 7, No. 1

Isolation of a *Saccharomyces cerevisiae* Centromere DNA-Binding Protein, Its Human Homolog, and Its Possible Role as a Transcription Factor

RICHARD J. BRAM AND ROGER D. KORNBERG*

Department of Cell Biology, Stanford University School of Medicine, Stanford, California 94305

Received 7 April 1986/Accepted 22 September 1986

A protein that binds specifically to *Saccharomyces cerevisiae* centromere DNA element I was purified on the basis of a nitrocellulose filter-binding assay. This protein, termed centromere-binding protein 1 (CP1), was heat stable and renaturable from sodium dodecyl sulfate (SDS), and assays of eluates from SDS gels indicated a molecular weight of 57,000 to 64,000. An activity with similar specificity and stability was detected in human lymphocyte extracts, and analysis in SDS gels revealed a molecular weight of 39,000 to 49,000. CP1-binding sites occurred not only at centromeres but also near many transcription units, for example, adjacent to binding sites for the *GAL4*-positive regulatory protein upstream of the *GAL2* gene in *S. cerevisiae* and adjacent to the TATA element of the adenovirus major late promoter. A factor (termed USF) that binds to the latter site and stimulates transcription has been isolated from HeLa cells by others.

Centromeres are specialized regions of eucaryotic chromosomes that form sites of attachment for spindle microtubules in mitosis and meiosis (13, 35, 36). Little is known of the mechanism of spindle attachment or the molecules involved. We report here on the isolation of a centromere DNA-binding protein from *Saccharomyces cerevisiae* and its counterpart from human cells.

Cloning and deletion analysis of yeast centromeres revealed three contiguous conserved sequences essential to or important for centromere function: centromere DNA element I (CDEI), an 8-base-pair (bp) conserved sequence; CDEII, an apparently random, adenine-plus-thymine (A + T)-rich stretch of approximately 90 bp; and CDEIII, a 25-bp conserved sequence (11, 12, 17, 18, 20, 21, 24, 29, 32, 33, 39). Nuclease digestion studies of chromatin containing centromere regions *CEN3* and *CEN11* revealed an unusual configuration: 250 bp, including the three conserved DNA elements, are especially protected from digestion (4). Crude fractions from *S. cerevisiae* bind naked DNA containing centromere sequences (5), but no resolution of the proteins involved has been described.

A centromere-specific DNA-binding activity was discovered during recent studies of the *GAL4*-positive regulatory protein (7). DNase I footprints with a yeast nuclear extract showed three 30-bp regions of protection upstream of the *GAL2* gene: two regions located 55 bp apart, and the third located 25 bp further upstream. The first two regions were attributed to *GAL4* protein binding, and the third was attributed to another activity for the following reasons. (i) Protection of the first two regions was abolished by competition with a *GAL4*-binding oligonucleotide, while protection of the third region was unaffected. (ii) The first two regions were homologous with a consensus *GAL4*-binding sequence, while the third region was altogether different, showing a striking homology to many CDEIs. During studies of protein binding to the third region (see below), a further homology came to light, suggesting a relationship to transcription.

* Corresponding author.

MATERIALS AND METHODS

Plasmid DNAs. Plasmid pG2p contains the 5′ half of the *GAL2* gene (7). Plasmids p181, p113, p200, p201, p203, p202, and p182 containing *Hin*dIII fragments of *CEN1*, *CEN4*, *CEN7*, *CEN11*, *CEN14*, *CEN15*, and 2μm-STB in *YRp14/ARS1* were kindly provided by P. Hieter (20).

Yeast strains and media. BY2 (6) is a protease-deficient strain (*pep4-3*) carrying the *GAL4* gene on a high-copy-number plasmid. YNN 267 (**a** *ura3-52 Δhis3-200 ade2-101 lys2-801 met⁻ Δgal4-537*) was a gift from M. Johnston. BJ926 (α *trp1 prc1-126 pep4-3 prb1-1122 can1*/**a** *his1 prc1-126 pep4-3 prb1-1122 can1*) was kindly provided by E. Jones. Cells were grown in YP medium (1% yeast extract, 2% Bacto-Peptone) containing 2% glucose to an A_{600} of 5.

Nitrocellulose filter-binding assay and DNase I footprints. Plasmid pG2p DNA was linearized by cleavage with *Eco*RI, and the recessed 3′ ends were filled in with the large fragment of DNA polymerase I in the presence of [α-^{32}P]dATP and cold TTP. In some experiments, a 0.28-kilobase (kb) fragment containing the two *GAL4*-binding sites and the centromere-binding protein 1 (CP1)-binding site prepared by cleavage of pG2p with *Hin*dIII and *Hpa*I and agarose gel electrophoresis was used for greater sensitivity. A palindromic oligonucleotide (Fig. 1) was synthesized, purified, and self-annealed (6). Details of the nitrocellulose filter-binding and DNase I footprint experiments were previously described (6). Other probes for filter binding were prepared by excision of CEN-containing fragments with *Hin*dIII, labeling, and isolation as described above. Probes for DNase I footprint experiments were prepared by cleavage with *Bgl*II (*CEN7* and *CEN15*) or *Hin*dIII (*CEN11*) after labeling as described above. Polymerase I was inactivated by heating at 65°C for 20 min. The labeled DNAs were cleaved with *Xba*I (*CEN15* and *CEN11*) or *Hin*dIII (*CEN7*), and CEN-containing fragments were isolated by agarose gel electrophoresis.

Yeast cell extracts and protein fractionation. Crude extracts and nuclear extracts were prepared as described previously

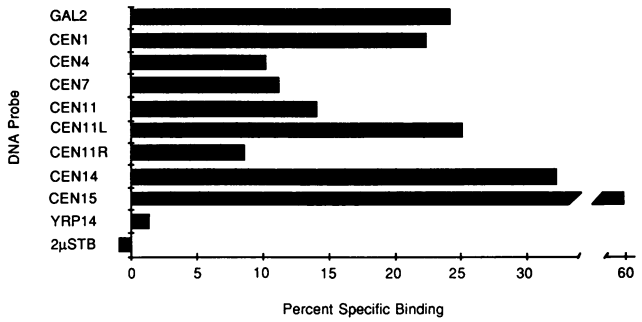

FIG. 1. Specific protein binding to yeast centromere DNAs. ^{32}P-labeled DNA fragments containing centromere, 2μm-STB, or vector (*YRP14/ARS1*) sequences were treated with crude extract from strain YNN267 (5 μg of protein) or with phosphocellulose fractions in the case of *CEN15* (0.6 μg of protein) and filtered through nitrocellulose. Percent specific binding is ^{32}P bound to the filter in the absence of a competing, synthetic oligonucleotide (Table 1) minus that bound in its presence, all divided by ^{32}P bound in the absence of the oligonucleotide and multiplied by 100%. ^{32}P bound in the absence of oligonucleotide was at least 1,400 cpm in all cases.

(6, 7). BJ926 cells were harvested by centrifugation, washed with cold water, and suspended with an amount of 10 mM Tris (pH 7.5)–1.5 mM $MgCl_2$–15 mM KCl–0.1 mM EDTA–1 mM phenylmethylsulfonyl fluoride–2 μM pepstatin A–0.6 μM leupeptin–0.5 mM dithiothreitol (9) equal to the volume of the cell pellet. The suspension was treated with an equal volume of glass beads (diameter, 0.45 mm) and eight 30-s pulses of a bead beater (Biospec Products, Bartlesville, Okla.). The cell lysate was centrifuged in a Sorvall SS34 rotor at 10,000 rpm for 10 min at 4°C, and the supernatant was made isotonic by addition of 0.1 volume of 250 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (pH 7.5)–50 mM $MgCl_2$–1 mM EDTA–5 mM dithiothreitol–500 mM KCl. The mixture was centrifuged in a Beckman Ti60 rotor at 55,000 rpm for 3 h at 0°C, and the supernatant (14 ml with 9.5 mg of protein per ml) was loaded at 0.7 ml/min on a 30-ml phosphocellulose column (diameter, 2.5 cm) equilibrated with buffer A50 as described previously (6). The column was washed with 80 ml each of buffer A50 and buffer A250, and CP1 activity was eluted with 50 ml of buffer A400. Peak fractions were pooled, adjusted to 70% of saturation with solid ammonium sulfate, and centrifuged. The precipitate was resuspended in buffer A50, dialyzed twice against 100 volumes of buffer A50 for 3 h at 4°C, and stored at −80°C.

Human lymphocyte extract. MICH cells (an Epstein-Barr virus-immortalized human B lymphocyte line) were kindly provided by J. Ways and D. Goldfarb. Cells were grown in suspension in 200 ml of RPMI 1640 plus 5% fetal bovine serum at 37°C to a density of 5×10^5 to 10×10^5/ml, harvested by centrifugation, washed wth 10 ml of buffer A50, and lysed by adding 2.7 ml (9 pellet volumes) of distilled water, followed by 20 strokes of a Dounce homogenizer. Buffer A500 (0.1 volume) was added, and nuclei were collected by centrifugation at 3,000 rpm for 10 min at 0°C. The pellet was suspended in 2.7 ml of buffer A50, and solid ammonium sulfate was added to 0.3 M. After gentle mixing at 4°C for 20 min, the solution was clarified by centrifugation

TABLE 1. Homology between DNA sequence upstream of the *GAL2* gene and sequences of yeast centromere DNA elements I[a]

Genetic element	Sequence	No. of matches/22 bp to:	
		CDEI consensus	Symmetrized *GAL2*
CEN1	AGTCTTGTCACATGACATAATA	17	13
CEN3	AAATAAGTCACATGATGATATT	20	16
CEN4	CAAAAGGTCACATGCTTATAAT	16	14
CEN6	CTTTTCATCACGTGCTATAAAA	17	14
CEN7	AAATATATCACGTGTTATATTT	20	19
CEN10	AACTTAATCACGTGTTAAATAA	19	15
CEN11	TCATAAGTCACATGATAAAAAC	19	14
CEN14	TAGTTAGTCACGTGCAGCTTTT	15	14
CEN15	AATAATATCACGTGAACTTATT	17	16
CEN16	AAATAGATCACATGATATATTT	21	20
GAL2	AAATGGGTCACGTGATCTATAT	19	22
CDEI consensus	AA$^{T}_{A}$T$^{T}_{A}$A$^{G}_{A}$TCAC$^{G}_{A}$TGATA$^{A}_{T}$A$^{A}_{T}$A$^{A}_{T}$T		
Symmetrized *GAL2*	A$^{T}_{A}$AT$^{A}_{G}$G$^{A}_{G}$TCACGTGA$^{C}_{T}$C$^{C}_{T}$AT$^{T}_{A}$T		
Synthetic oligonucleotide	AATTAAAATAGATCACGTGATCTATTTT		

[a] *GAL2* sequence is from Bram et al. (7). Centromere sequences are from Hieter et al. (20) for *CEN1*, *7*, *10*, *14*, *15*, and *16*; Fitzgerald-Hayes et al. (18) for *CEN3* and *11*; and Panzeri and Philippsen (33) for *CEN6*. CDEI was defined by Hieter et al. (20) as the sequence Pu-T-C-A-C-Pu-T-G near the center of the consensus shown above. Matches to the symmetrized *GAL2* sequence are overlined.

TABLE 2. Partial list of sequences homologous to the *GAL2*-associated CP1-binding site

Genetic element	Sequence[a]	Location or source (reference)
GAL2	AA(T/A)AT(A/G)G(A/G)TCACGTGA(C/T)C(C/T)AT(T/A)TT	(7)
Yeast *CEN11-MET14*	CACTAATTTCACGTGATC	510 bp 3′ of *CDEIII* (18)
Yeast *CBP2*	ACACGCCCTCACGTGAGTGAATTT	285 bp 5′ of AUG (25)
Yeast Arg-tRNA 3a	AGTTACCATCACGTGCCGCTCTAT	475 bp 5′ of tRNA start (42)
Yeast *TRP1-ARS1*	ACTATTGAGCACGTGAGTATACGT	51 bp 5′ of *TRP1* AUG (41)
Yeast *ADR2*	GGAATGTTCCACGTGAAGCTATCT	734 bp 5′ of mRNA start (43)
Bovine 1.706 satellite DNA	TGCACTGATCACGTGACTGATCAT	23 bp prototype (34)
Adenovirus major late promoter	GGTGTAGGCCACGTGACCGGGTGT	60 bp 5′ of mRNA start (37)
Mouse middle repetitive LTR[b]-like sequence	ATAAAGTCTCACGTGGTTTGCAAC	No. 38 of Wirth et al. (44)
Human *Kpn*I repeat	CAATGAGATCACGTGGACACAGGA	No. 52 of Digiovanni et al. (15)
Drosophila melanogaster transposable element *MDG3*	ATACAATTTCACGTGTCTCTTTTA	No. 111 at 3′ end, no. 78 at 5′ end of Bayev et al. (3)
D. melanogaster transposable P	AAATTAATTCACGTGCCGAAGTGT	No. 195 of O'Hare and Rubin (31)
Trypanosome kinetoplast DNA minicircle	TACACAAATCACGTGCTATTTTCG	No. 746 of Barrois et al. (2)
Xenopus laevis histone H2A1	ATCAGAGCTCACGTGATCACATGG	71 bp 3′ of stop codon 200–300 bp 3′ of *H4* gene (27)
Human *Alu* sequences 5′ of globin genes	CAGGCAGATCACCTGAGGTCAAGA	55 in *Alu* upstream of Gγ and δ (16)
Alu sequence 5′ of human ε-globin gene	TGGGCAGATCATGAGGTCAAGAGA	in *Alu* sequence 1.5 kb 5′ of ε AUG (1)
Gross passage A murine leukemia virus LTR	ATTGTGAATCACGTGAATAAAAGA	92 bp from 3′ end of provirus (14)
Mouse adenovirus strain F1 inverted terminal repeat	CCCGGGTTTCACGTGGTGCGTCAG	No. 88 of Temple et al. (40)
HTLVIII LTRs	GCATTTCATCACGTGGCCCGAGAG	176 bp 5′ of mRNA start also 365 bp 5′ of 3′ end (28)
Bovine papilloma virus	CCTCTAAATCACGTGGCATTTTAA	No. 4922 of Chen et al. (10)

[a] Matches to the symmetrized *GAL2*-associated sequence are overlined.
[b] LTR, Long terminal repeat.

in a Sorvall AH650 rotor at 50,000 rpm for 1 h at 0°C. The supernatant was frozen in liquid N_2 and stored at −80°C.

Recovery of CP1 activity from SDS-polyacrylamide gels. Partially purified yeast CP1 (60 μg of peak phosphocellulose fractions) or crude extracts containing human CP1 (270 μg after dialysis against buffer A50) were subjected to electrophoresis in 1.5-mm-thick sodium dodecyl sulfate (SDS)–10% polyacrylamide gels (23) calibrated with SDS-polyacrylamide gel electrophoresis low-molecular-weight standards (Bio-Rad Laboratories, Richmond, Calif.). Proteins were eluted from gel slices as described before (19), precipitated, and renatured by resuspension in 5 μl of 6 M guanidine hydrochloride–30 mM Tris (pH 8), followed by dilution with 250 μl of 50 mM Tris (pH 7.5)–0.1 mM EDTA–1 mM dithiothreitol–0.1 mg of bovine serum albumin per ml–0.15 M NaCl–20% glycerol. Approximately 75% of the yeast and >50% of the human CP1 activity was recovered after incubation of the diluted protein at 22°C for 1 h.

RESULTS

Homology of a protein-binding sequence upstream of the *GAL2* gene to CDEI. The protein-binding sequence located upstream of the *GAL2* gene, 25 bp beyond a pair of *GAL4*-binding sites, was dyad symmetric at 18 of 24 bp, with the central 8 bp forming the perfect palindrome T-C-A-C-G-T-G-A. This *GAL2*-associated sequence was homologous at 19 of 22 bp with a consensus sequence (Table 1) of 10 CDEIs (compared with 15 to 21 correct matches to the consensus of the CDEI sequences themselves). Comparison with the *GAL2*-associated sequence revealed a previously unrecognized dyad symmetry of CDEI sequences (with 20 of 22 bp of dyad-symmetric sequence in the CDEI of *CEN7*).

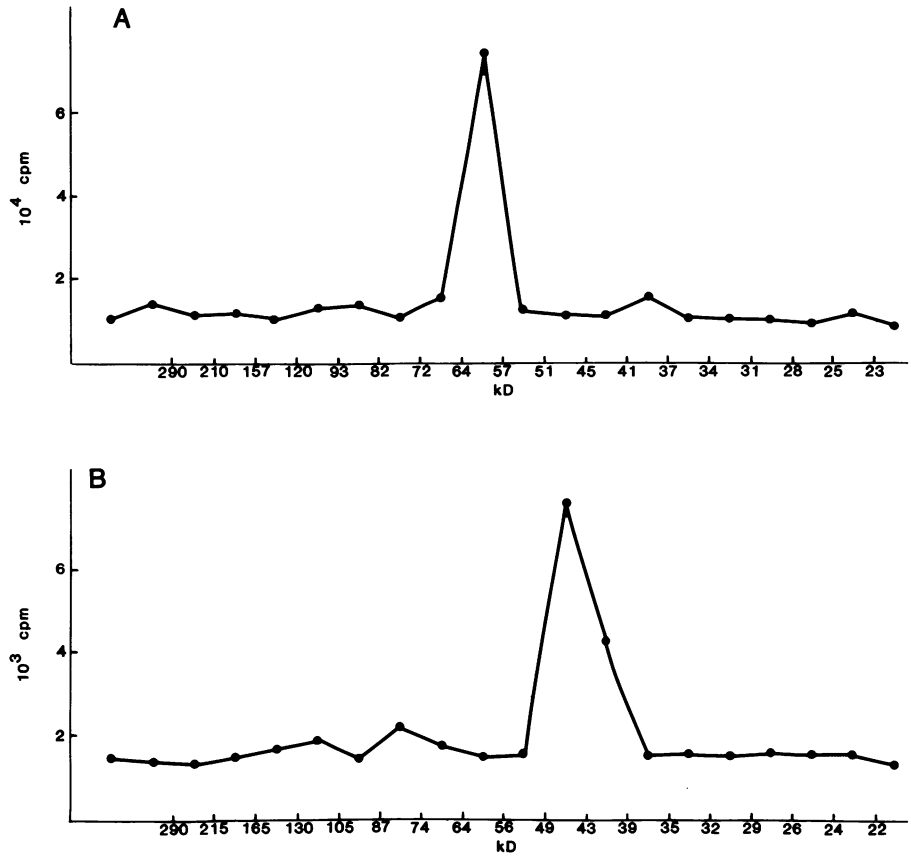

FIG. 2. Distribution of CEN-binding activity in SDS gels of (A) yeast phosphocellulose fractions and (B) human lymphocyte nuclear extract. ^{32}P-labeled pG2p DNA was incubated with 5 μl of renatured protein from each gel slice and filtered through nitrocellulose as described. ^{32}P counts per minute bound to filters are plotted for slices from the top of the gel at the left. Major peaks of binding activity in panels A and B were abolished by adding a competing, synthetic oligonucleotide.

Specific protein binding to CEN sequences. Protein binding to *GAL2*-associated and CEN sequences was detected in nitrocellulose filter-binding assays. ^{32}P-labeled plasmid DNAs (1 to 2 ng) containing these sequences were mixed with protein fractions and carrier DNA (1 to 2 μg) and passed through nitrocellulose. Binding specific for the *GAL2*-associated sequence was determined from the difference between label bound in the presence and absence of a 50-fold molar excess (4 ng) of synthetic oligonucleotide containing the *GAL2* sequence. In a typical experiment with labeled plasmid containing the *GAL2* upstream region and crude yeast extract (5 μg of protein), 35% of the label bound in the presence, and 65% bound in the absence of competing oligonucleotide. All six CEN DNAs tested showed binding specific for the *GAL2*-associated sequence, whereas DNAs containing chromosome replication or stabilization elements (*ARS1*, *ARS14*, and 2μm-STB, a *cis*-acting locus required for stable maintenance of the 2μm plasmid) showed no specific binding (Fig. 1). The variation in the percentage of specific binding among the CEN DNAs was due to variation in the length and sequence composition of the restriction fragments used. This is illustrated for a 2.9-kb *Hin*dIII fragment containing *CEN11* and 0.8-kb (*CEN11L*, containing *CEN11*) and 2.1-kb (*CEN11R*) products of cleaving this fragment with *Xba*I, which showed 14.1, 25.1, and 8.6% specific binding, respectively (Fig. 1). The increase in percent specific binding upon shortening the *CEN11*-containing fragment may be attributed to the removal of nonspecific binding to flanking DNA. The effect would have been greater, were it not for some specific binding to the flanking DNA (*CEN11R*), possibly to a sequence between *CEN11* and the *MET14* gene identical with the central eight residues of the *GAL2*-associated sequence (Table 2).

Purification and properties of CEN-binding protein. CEN-binding activity was heat stable and renaturable from SDS (Heating a crude extract for 10 min at 100°C resulted in a 23% loss of activity and in the precipitation of 98% of the total protein; removal of SDS by acetone precipitation and renaturation from guanidine hydrochloride resulted in 76% recovery of activity). CEN-binding activity was also insensitive to RNase treatment (36% loss of activity upon treatment with 1 μg of RNase A per μl for 40 min at 37°C), but it was completely destroyed by proteinase K.

Crude extracts were fractionated on phosphocellulose (160-fold enrichment of CEN-binding activity in 54% yield) or they were heated, centrifuged, and fractionated on Affi-Gel Blue (875-fold enrichment of CEN-binding activity in 22% yield). Peak phosphocellulose or Affi-Gel Blue fractions were subjected to SDS-polyacrylamide gel electrophoresis, and eluates of gel slices were treated with acetone and guanidine hydrochloride for renaturation of CEN-binding activity. Nitrocellulose filter-binding assays revealed a single peak of activity in the 57- to 64-kilodalton region of the gel (Fig. 2A).

DNAse I footprint of CEN-binding protein on CEN DNAs. End-labeled DNA fragments containing CEN sequences

were mixed with peak phosphocellulose fractions, digested briefly with DNase I, and subjected to polyacrylamide gel electrophoresis and autoradiography. *CEN7*, *CEN11* (Fig. 3) and *CEN15* (not shown) showed protected regions of about 25 bp centered on CDEI but no protection of CDEs II and III. The size and location of the protected regions were similar to those previously found for the *GAL2*-associated sequence (7).

CEN-binding activity in human cells. Fitzgerald-Hayes et al. (17) noted a homology of 12 of 14 bp between CDEI and the bovine 1.706 satellite DNA (a 23-bp repeating structure). The region of homology included the sequences involved in the CEN-binding interaction described above. The *GAL2*-associated binding sequence was identical with the prototype bovine sequence (Table 2) at 11 of 11 central bp. This prompted us to investigate the occurrence of CEN-binding activity in mammalian cells. The nitrocellulose filter-binding assay described above was applied to extracts from human lymphocyte nuclei with a labeled *Hin*dIII-*Hpa*I DNA fragment containing the *GAL2* upstream region. Significant DNA-binding activity specific for the *GAL2*-associated sequence was detected (31% of label bound in the absence, and 6% bound in the presence of oligonucleotide containing the *GAL2*-associated sequence with 5 μg of protein; oligonucleotides containing unrelated sequences had no effect). The activity was heat stable (43% recovery of activity after treatment at 100°C for 10 min) and renaturable from SDS. Upon electrophoresis in SDS-polyacrylamide gels and elution from gel slices, a single peak of activity in the 39- to 49-kilodalton region of the gel was obtained (Fig. 2B).

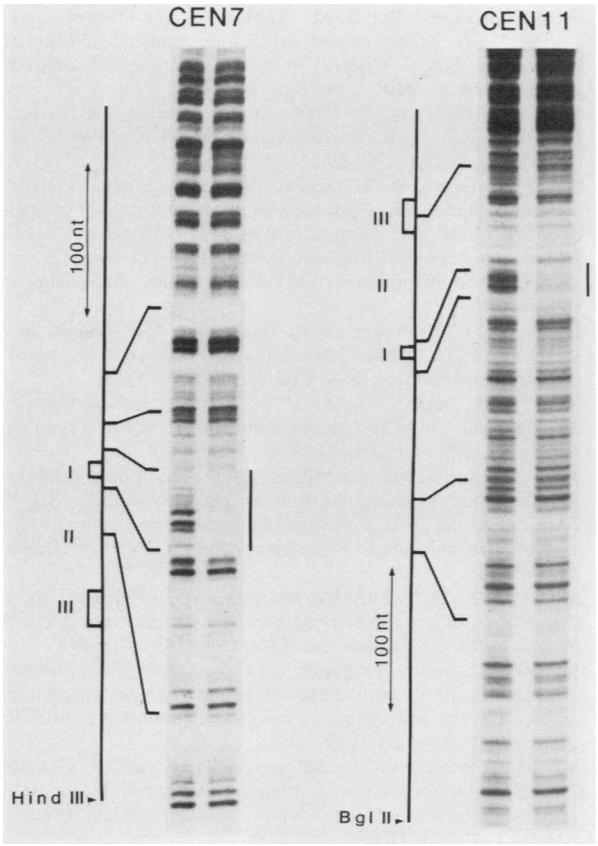

FIG. 3. DNase I footprints of CEN-binding activity on *CEN7* and *CEN11* DNAs. ^{32}P-labeled CEN DNA fragments were incubated with (right lanes) or without (left lanes) peak phosphocellulose fractions (12 μg of protein) and processed as described in the text. The correspondence between points in the CEN DNA maps and positions of DNA bands was determined from patterns of φX174/*Hae*III fragments (not shown). Abbreviations: nt, nucleotides; I, CDEI; II, CDEII; III, CDEIII.

DISCUSSION

Although the polypeptides from yeast responsible for binding six CEN sequences and a *GAL2*-associated sequence were not purified to homogeneity, several lines of evidence suggested that the various binding activities were identical. First, the various DNA sequences were highly homologous, and an oligonucleotide containing one sequence competed for binding with all the others. Second, DNase I footprints showed that the location of protein binding to all the sequences was the same. Finally, the binding activity migrated as a single band in an SDS-polyacrylamide gel. We refer to what is, in all likelihood, a single DNA-binding polypeptide as CP1.

The striking features of CP1 are the occurrence of noncentromeric binding sites and the apparent conservation of the protein through evolution. In addition to a *GAL2*-associated site, a computer-assisted search of a DNA sequence data bank revealed several possible binding sequences in yeast (Table 2). These sequences were homologous at 8 or more of the 12 residues containing the central palindrome C-A-C-G-T-G and were located upstream of various transcription units. The abundance of CP1 in *S. cerevisiae* was in keeping with its action at sites in addition to the few associated with centromeres. The amount of DNA specifically bound by crude extracts led to an estimate of approximately 500 copies of CP1 per cell. This is probably a low estimate that does not take into account the protein associated with cellular DNA, since the extracts were prepared under conditions similar to those used to measure binding, and the complex was quite stable under these conditions (half-life for dissociation, ca. 20 min).

Yeast CP1 and the CP1 activity identified in human lymphocyte extracts shared both specificity for CDEI sequences and unusual stability and capacity for renaturation. These findings, together with the homology between yeast CDEIs and a bovine satellite known to occur in centromeres (22, 38), lead us to suggest that the yeast and human proteins perform similar functions. An indication of a possible function that would fit with the occurrence of noncentromeric sites comes from recent reports (8, 26, 37) of a transcription factor from HeLa cells, termed USF, whose recognition sequence in adenovirus DNA is identical at 13 out of 14 central bp with the *GAL2*-associated CP1-binding sequence (Table 2). USF not only recognizes the same DNA sequence as CP1, but is heat stable as well, and it seems likely that USF is analogous to CP1 in lymphocyte extracts. USF stimulates transcription from the adenovirus major late promoter 10- to 20-fold in vitro, and its binding appears to be cooperative with that of a factor that recognizes the TATA sequence. CP1 may play a similar role in *S. cerevisiae*, assisting the binding of *GAL4* protein, microtubule proteins, and the like. Preliminary experiments indeed suggested that CP1 facilitates *GAL4*-binding to the *GAL2* upstream region (Bram, unpublished data). The occurrence of a CP1-binding site in this region but not upstream of other galactose-inducible genes may be explained by the low level of *GAL4* protein in cells and the advantage of transcribing the *GAL2* gene, which encodes a galactose permease, before the other *GAL* genes.

Further consistent with a role of CP1 in facilitating the binding of other proteins, it appears to be important for but not essential for centromere function. Deletion of CDEI from *CEN6* resulted in partial loss of mitotic stability (about 50% retention of a *CEN6/ARS1/TRP1* plasmid after 10 generations) and no impairment of segregation in meiosis (32). Other functions for CP1 besides an accessory role in microtubule attachment could also account for these observations. For example, CP1-binding to CDEI may prevent processes such as transcription from interfering with centromere activity (Panzeri et al. [32] mention a deleterious effect of transcription on *CEN6* activity). CP1 binding could play a similar role in the case of the *GAL2* gene, sequestering the *GAL4*-binding sites, protecting them from transcription originating upstream, or preventing their action on genes located upstream. With the latter possibility in mind, we placed a restriction fragment from the *GAL2* regulatory region containing the CP1 and *GAL4*-binding sites in front of a *GAL1-HIS3* fusion gene. When the orientation of the restriction fragment was inverted (CP1-binding site between the *GAL4*-binding sites and the promoter), the level of *GAL1-HIS3* transcription was only 20% of that found with the fragment in the normal orientation, indicative of some capacity of CP1 to block *GAL4* action. The persistence of appreciable transcription, however, argues against such a function for the primary role of CP1 in vivo.

The human homolog of CP1 is probably distinct from a centromeric antigen recognized by autoantibodies in some CREST-scleroderma patients (30). CP1 activity is released from lymphocyte nuclei with 0.3 M ammonium sulfate (I = 0.9) and migrates in SDS-polyacrylamide gels as a polypeptide of 39 to 49 kilodaltons, whereas CREST antigen is extracted with 4 M NaCl and migrates in gels as a 70-kilodalton species.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health grants to R.D.K. R.J.B. is the Eva S. Gutmann Scholar in the Department of Cell Biology, and received support from the Medical Scientist Training Program grant GM07365 from the National Institute of General Medical Sciences.

LITERATURE CITED

1. **Baralle, F. E., C. C. Shoulders, S. Goodbourn, A. Jeffreys, and N. Proudfoot.** 1980. The 5′ flanking region of human ε-globin gene. Nucleic Acids Res. **8**:4393–4404.
2. **Barrois, M., G. Riou, and F. Galibert.** 1981. Complete nucleotide sequence of minicircle kinetoplast DNA from *Trypanosoma equiperdum*. Proc. Natl. Acad. Sci. USA **78**:3323–3327.
3. **Bayev, A. A., Jr., A. S. Krayev, N. V. Lyubomirskaya, Y. V. Ilyin, K. G. Skryabin, and G. P. Georgiev.** 1980. The transposable element MDG3 in *Drosophila melanogaster* is flanked with perfect direct and mismatched inverted repeats. Nucleic Acids Res. **8**:3263–3273.
4. **Bloom, K., and J. Carbon.** 1982. Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. Cell **29**:305–317.
5. **Bloom, K. S., M. Fitzgerald-Hayes, and J. Carbon.** 1983. Structural analysis and sequence organization of yeast centromeres. Cold Spring Harbor Symp. Quant. Biol. **157**:1175–1185.
6. **Bram, R. J., and R. D. Kornberg.** 1985. Specific protein binding to far upstream activating sequences in polymerase II promoters. Proc. Natl. Acad. Sci. USA **82**:43–47.
7. **Bram, R. J., N. F. Lue, and R. D. Kornberg.** 1986. A *GAL* family of upstream activating sequences in yeast: roles in both induction and repression of transcription. EMBO J. **5**:603–608.
8. **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.
9. **Celniker, S. E., and J. L. Campbell.** 1982. Yeast DNA replication in vitro: initiation and elongation events mimic in vivo processes. Cell **31**:201–213.
10. **Chen, E. Y., P. M. Howley, A. D. Levinson, and P. H. Seeburg.** 1982. The primary structure and genetic organization of the bovine papillomavirus type 1 genome. Nature (London) **299**:529–534.
11. **Clarke, J., and J. Carbon.** 1980. Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature (London) **287**:504–509.
12. **Clarke, L., and J. Carbon.** 1983. Genomic substitutions of centromeres in *Saccharomyces cerevisiae*. Nature (London) **305**:23–28.
13. **Comings, D. E., and T. A. Okada.** 1971. Fine structure of kinetochora in Indian muntjac. Exp. Cell Res. **67**:97–110.
14. **DesGroseillers, L., R. Villemur, and P. Jolicoeur.** 1983. The high leukemogenic potential of gross passage A murine leukemia virus maps in the region of the genome corresponding to the long terminal repeat and to the 3′ end of *env*. J. Virol. **47**:24–32.
15. **Digiovanni, L., S. R. Haynes, R. Misra, and W. R. Jelinek.** 1983. KpnI family of long-dispersed repeated DNA sequences of man: evidence for entry into genomic DNA of DNA copies of polyA terminated KpnI RNAs. Proc. Natl. Acad. Sci. USA **80**: 6533–6537.
16. **Duncan, C. H., P. Jagadeeswaran, R. R. C. Wang, and S. M. Weissman.** 1981. Structural analysis of templates and RNA polymerase III transcripts of *Alu* family sequences interspersed among the human β-like globin genes. Gene **13**:185–196.
17. **Fitzgerald-Hayes, M., J.-M. Buhler, T. G. Cooper, and J. Carbon.** 1982. Isolation and subcloning analysis of functional centromere DNA (*CEN11*) from *Saccharomyces cerevisiae* chromosome XI. Mol. Cell. Biol. **2**:82–87.
18. **Fitzgerald-Hayes, M., L. Clarke, and J. Carbon.** 1982. Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. Cell **29**:235–244.
19. **Hager, D. A., and R. R. Burgess.** 1980. Elution of proteins from sodium dodecyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity: results with sigma subunit of *Escherichia coli* RNA polymerase, wheat germ DNA topoisomerase, and other enzymes. Anal. Biochem. **109**:76–86.
20. **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.
21. **Hsiao, C.-L., and J. Carbon.** 1981. Direct selection procedure for the isolation of functional centromeric DNA. Proc. Natl. Acad. Sci. USA **78**:3760–3764.
22. **John, B., and G. L. C. Miklos.** 1979. Functional aspects of satellite DNA and heterochromatin. Int. Rev. Cytol. **58**:1–113.
23. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
24. **Maine, G. T., R. T. Surosky, and B.-K. Tye.** 1984. Isolation and characterization of the centromere from chromosome V (*CEN5*) of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:86–91.
25. **McGraw, P., and A. Tzagoloff.** 1983. Assembly of the mitochondrial membrane system: characterization of a yeast nuclear gene involved in the processing of the cytochrome *b* pre-mRNA. J. Biol. Chem. **258**:9459–9468.
26. **Miyamoto, N. G., V. Moncollin, J. M. Egly, and P. Chambon.** 1985. Specific interaction between a transcription factor and the upstream element of the adenovirus-2 major late promoter. EMBO J. **4**:3563–3570.
27. **Moorman, A. F. M., P. A. J. DeBoer, R. T. M. DeLaaf, and O. H. J. Destree.** 1982. Primary structure of the histone H2A and H2B genes and their flanking sequences in a minor histone gene cluster of *Xenopus laevis*. FEBS Lett. **144**:235–241.
28. **Muesing, M. A., D. H. Smith, C. C. Cabradilla, C. V. Benton, L. A. Lasky, and D. J. Capon.** 1985. Nucleic acid structure, and expression of the human AIDS/lymphadenopathy retrovirus. Nature (London) **313**:450–458.

29. **Neitz, M., and J. Carbon.** 1985. Identification and characterization of the centromere from chromosome XIV in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **5**:2887–2893.
30. **Nishikai, M., Y. Okano, H. Yamashita, and M. Watanabe.** 1984. Characterization of centromere (kinetochore) antigen reactive with sera of patients with a scleroderma variant (CREST syndrome). Ann. Rheum. Dis. **43**:819–824.
31. **O'Hare, K., and G. M. Rubin.** 1983. Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. Cell **34**:25–35.
32. **Panzeri, L., L. Landonio, A. Stotz, and P. Philippsen.** 1985. Role of conserved sequence elements in yeast centromere DNA. EMBO J. **4**:1867–1874.
33. **Panzeri, L., and P. Philippsen.** 1982. Centromeric DNA from chromosome VI in *Saccharomyces cerevisiae*. EMBO J. **1**: 1605–1611.
34. **Pech, M., R. E. Streeck, and H. G. Zachau.** 1979. Patchwork structure of a bovine satellite DNA. Cell **18**:883–893.
35. **Ris, H., and P. L. Witt.** 1981. Structure of the mammalian kinetochore. Chromosoma **82**:153–170.
36. **Roos, U. P.** 1973. Light and election microscopy of rat kangaroo cells in mitosis. Chromosoma **41**:195–200.
37. **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.
38. **Singer, M.** 1982. Highly repeated sequences in mammalian genomes. Int. Rev. Cytol. **76**:67–112.
39. **Stinchcomb, D. T., C. Mann, and R. W. L. Davis.** 1982. Centromeric DNA from *Saccharomyces cerevisiae*. J. Mol. Biol. **158**:157–179.
40. **Temple, M., G. Antoine, H. Delius, S. Stahl, and E.-L. Winnacker.** 1981. Replication of mouse adenovirus strain F1 DNA. Virology **109**:1–12.
41. **Tschumper, G., and J. Carbon.** 1980. Sequence of a yeast DNA fragment containing a chromosomal replicator and the TRP1 gene. Gene **10**:157–166.
42. **Villanueva, J., P. Bull, P. Valenzuela, and A. Vanegas.** 1984. Nucleotide sequence of a yeast Arg-tRNA-3A gene and its transcription in a homologous in vitro system. FEBS Lett. **167**:165–169.
43. **Williamson, V. M., D. Cox, E. T. Young, D. W. Russell, and M. Smith.** 1983. Characterization of transposable element-associated mutations that alter yeast alcohol dehydrogenase II expression. Mol. Cell. Biol. **3**:20–31.
44. **Wirth, T., K. Gloeggler, T. Baumruker, M. Schmidt, and I. Horuk.** 1983. Family of middle repetitive DNA sequences in the mouse genome with structural features of solitary retroviral long terminal repeats. Proc. Natl. Acad. Sci. USA **80**:3327–3330.