Identification of a Subdomain of CENP-B That Is Necessary and Sufficient for Localization to the Human Centromere

Ann F. Pluta, Noriko Saitoh, Ilya Goldberg, and William C. Earnshaw

Department of Cell Biology and Anatomy, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Abstract. We have combined in vivo and in vitro approaches to investigate the function of CENP-B, a major protein of human centromeric heterochromatin. Expression of epitope-tagged deletion derivatives of CENP-B in HeLa cells revealed that a single domain <158 residues from the amino terminus of the protein is sufficient to localize CENP-B to centromeres. Centromere localization was abolished if as few as 28 amino acids were removed from the amino terminus of CENP-B. The centromere localization signal of CENP-B can function in an autonomous fashion, relocating a fused bacterial enzyme to centromeres. The centromere localization domain of CENP-B specifically binds in vitro to a subset of α-satellite DNA monomers. These results suggest that the primary mechanism for localization of CENP-B to centromeres involves the recognition of a DNA sequence found at centromeres. Analysis of the distribution of this sequence in α-satellite DNA suggests that CENP-B binding may have profound effects on chromatin structure at centromeres.

THE centromere (primary constriction) of human mitotic chromosomes consists of a compacted network of heterochromatin containing highly repeated satellite DNA sequences and their associated proteins. The kinetochore, the site of microtubule attachment to the chromosomes (Luykx, 1965; Brinkley and Stubblefield, 1966; Jokelainen, 1967), is situated at the surface of the centromeric heterochromatin. The kinetochore is now believed to be the location of the mechanochemical motors (such as dynein) (Pfarr et al., 1990; Steuer et al., 1990; Wordeman et al., 1991) that move chromosomes in mitosis. Both gene expression and recombination are strongly suppressed in regions of heterochromatin (Brown, 1966; Henikoff, 1981; Daniels et al., 1986). The interrelationship between the kinetochore and the underlying heterochromatin is unknown.

Centromeric heterochromatin is largely composed of α-satellite DNA, a highly polymorphic family of repetitive DNA sequences that comprises ∼5% of the human genome (Willard and Waye, 1987). The role of α-satellite DNA in human chromosome structure and function is not known. The consensus monomer for α-satellite DNA (171 bp in length) is typically found in higher-order repeats that may be as large as 35-mers (Choo et al., 1991). Within a repeat, adjacent monomers commonly show up to 40% sequence divergence (Willard and Waye, 1987). Arrays of α-satellite DNA apparently span contiguous regions of the centromeric DNA of two to three megabases (Wevrick and Willard, 1989; Jabs et al., 1989).

In humans, only a single heterochromatin protein has been identified to date. This protein, CENP-B, was originally discovered when centromeric autoantigens (termed CENP antigens) recognized by sera of patients with scleroderma spectrum disease were identified by immunoblotting (Earnshaw

and Rothfield, 1985). CENP-B remains the best-characterized centromere protein (Earnshaw et al., 1987). Variability in staining of the different chromosomes by CENP-B–specific antibodies initially led to the suggestion that binding of the protein to centromeres might be determined by DNA sequence (Earnshaw et al., 1987). In confirmation of this, CENP-B was recently shown to bind specifically to a 17-bp sequence (termed the CENP-B box) present in some α-satellite DNA monomers, but not in others (Masumoto et al., 1989). CENP-B binds only to monomers containing the CENP-B box. The amount of CENP-B present at different centromeres varies because of differences in the sequence composition of the α-satellite DNA (Willard and Waye, 1987).

Recent studies from two laboratories using anticentromere autoantibodies have uncovered a link between the CENP antigens and kinetochore structure and function (Bernat et al., 1990, 1991). IgG purified from autoimmune patient sera inhibits mitotic events when injected into human cells (Simerly et al., 1990; Bernat et al., 1990). The antibodies affect two different aspects of kinetochore assembly, depending upon when during the cell cycle they gain entry into the nucleus. If introduced into the nucleus before the S/G2 transition, the antibodies interfere with the subsequent assembly of a functional outer kinetochore plate during mitosis (Bernat et al., 1991). As a result, the injected cells become arrested in a prometaphase-like state. The chromosomes bind microtubules, but apparently cannot move along them. If the antibodies enter the nucleus later, during G2, the cells assemble a fragile kinetochore that is disrupted when microtubules bind to it (Bernat et al., 1991). Chromosomes in these cells appear to complete prometaphase congression movements but the cells subsequently arrest in metaphase.

The distribution of CENP-B within the centromere provides a link between these alterations in the kinetochore and the centromeric heterochromatin. Immunoelectron microscopy studies show unequivocally that CENP-B is not detectable in the kinetochore itself (Cooke et al., 1990). Instead, the protein is distributed throughout the heterochromatin beneath the kinetochore. However, certain IgG preparations that affect kinetochore function and morphology after microinjection recognize only CENP-B in immunoblots, suggesting that one target of the antibody inhibition is likely to be CENP-B. Together, these results imply that the injected antibodies do not act directly at the kinetochore itself. Instead, it appears more plausible that assembly of the centromeric heterochromatin (where the CENP-B:α-satellite DNA complex is located) is essential for the subsequent morphogenesis of a functional kinetochore in mitosis.

The antibody inhibition results suggest that the assembly of CENP-B at the centromere is an early and important step in the pathway leading to the assembly of a functional kinetochore at mitosis. Here we study the first step in this process, which involves the localization of CENP-B to the centromere. Our studies reveal that this localization process involves the sequence-specific binding of CENP-B to the α-satellite DNA. This binding is directed by a single domain at the amino terminus of the protein. Analysis of the distribution of CENP-B boxes in α-satellite DNA suggests that the binding of CENP-B is likely to have profound effects on the structure of the centromeric heterochromatin.

Materials and Methods

Cell Culture and Transfections

HeLa cells were grown in DME (Gibco Laboratories, Grand Island, NY) with 5% calf serum (Hyclone Laboratories Inc., Logan, UT). For transfections, cells were plated onto 60-mm tissue culture dishes at ~80–90% confluency and given at least 5 h to attach to the substrate before transfection. (In some experiments, cells were plated at the same density directly onto sterile 18-mm² coverslips before transfection. Alternatively, cells were plated to coverslips after transfection.) Just before transfection, cells were washed twice with OPTIMEM (Gibco Laboratories) and transiently transfected with 10 μg plasmid DNA using Lipofectin Reagent (Bethesda Research Laboratories, Gaithersburg, MD) according to the manufacturer's instructions. Cells were exposed to the DNA for 12–24 h and then incubated in fresh complete media for up to an additional 24 h. All plasmids used for transfections were prepared by standard alkaline lysis procedures (Sambrook et al., 1989) followed by banding in CsCl density gradients and dialysis overnight at 4°C against sterile TE, pH 8.0.

DNA Constructs

Tagged Full-Length CENP-B and Carboxyl-Terminal Deletions. A 1,984-bp SmaI fragment, which contains the entire 599 amino acid open reading frame of CENP-B plus minimal upstream and downstream noncoding sequences, was isolated from a chimeric genomic/cDNA plasmid pG/CNPB7, reacted with BglII linkers and inserted into the unique BglII site of pECE/SKL (Ellis et al., 1986). After verifying that the CENP-B insert was in the correct orientation with respect to the SV-40 early promoter, the resulting plasmid, pB⁻9-599', was digested with BstEII (contained within the downstream noncoding portion of the CENP-B insert) and EcoRV (within the pECE polylinker), then self-ligated to generate pB⁻9-599. A 90-bp BamHI fragment encoding the E1 tag sequence was isolated from pSM224 (obtained from P. Chen and S. Michaelis, Johns Hopkins School of Medicine, Baltimore, MD) and inserted into the remaining (upstream) BglII site of pB⁻9-599. Sequence analysis and restriction mapping confirmed that the resulting construct, pCNPB1-599, contained two tandem copies of the E1 tag sequence, in correct orientation and reading frame, upstream of CENP-B coding sequences. All subsequent tagged constructs generated in

this study are based upon this original construct, and thus contain two tandem copies of the E1 tag followed by 9 amino acids encoded by CENP-B upstream genomic sequences followed by CENP-B and/or TRP E coding sequences. The large DNA fragment (~4.5 kb) produced by SstI digestion of pCNPB1-599 was isolated and self-ligated to generate pCNPB1-465. The large DNA fragment (~3.5 kb) produced by SstII digestion of pCNPB1-599 was isolated and self-ligated to generate pCNPB1-158.

TRPE Constructs

A ~1-kb DNA fragment encoding codons 2 through 323 of the bacterial TRPE gene product plus all of the pATH 11 polylinker was amplified from pATH 11 by the polymerase chain reaction using a 5' PCR primer, 5' TCCCCGCGGGCCCCCAAACACAAAAACCGAC 3' (overlapping SstII and ApaI sites underlined), and 3' PCR primer, 5'CTCATGTTTGACAG-CTT 3'. The resulting TRPE fragment was inserted into the HincII site of pUC 9, isolated from the resulting plasmid by double digestion with SstI and SstII, and inserted into the large fragment generated by SstI/SstII digestion of pCNPB1-599 to generate plasmid, pCNPB1-158TRPE. pCNPB1-158TRPE was digested with ApaI, which cuts within the third codon of CENP-B and just upstream of the TRPE coding sequence, and the resulting 4.1-kb DNA fragment was isolated and self-ligated to generate pTRPE.

Tagged Amino-Terminal Deletions

The 1,984-bp SmaI fragment containing the entire open reading frame of CENP-B isolated from pG/CNPB7 was inserted into the unique EcoRV site of pECE/SKL. After verifying that the CENP-B insert was in the correct orientation with respect to the SV-40 early promoter, the resulting plasmid, pA3 was digested with BglII and PvuI, and the large fragment was gel isolated and subjected to unidirectional ExoIII digestion. Deletion products were subsequently digested with EcoRI and BstXI to remove vector sequences, cloned into E1-tagged pECE which had been generated by digestion of pCNPB1-599 with ApaI, followed by blunt end formation with T4 DNA polymerase, and then digestion with EcoRI. Resulting tagged in-frame amino-terminal deletion derivatives of CENP-B were identified by restriction mapping followed by dideoxy sequence analysis from a oligonucleotide primer located just upstream of the E1 tags.

Alphoid DNA Subcloning and Generation of Biotinylated Monomers

Plasmid p17H8, containing a 2,712-bp alphoid DNA sequence from human chromosome 17 has been previously described (Waye and Willard, 1986). p17H8 was digested with BsmI (New England Biolabs, Beverly, MA) which cuts each of the 16 monomers near position 45 and/or near position 75, yielding a population of fragments representing monomers 2–15, which range in size from 138–200 bp. The fragments were purified from agarose gels (NA 45 paper), blunt ended with T4 DNA polymerase, and cloned into the SmaI site of pBluescript KS+ (Stratagene). DNA isolated from resulting transformants was sequenced to determine the specific monomers cloned. p17M15 contained the consensus CENP-B box on monomer 15 and p17M5 contained monomer 5 lacking the consensus sequence:

CENP-B box, consensus   PyTTCGTTGGAAPuCGGGA
CENP-B box, monomer 15   TTTCGTTGGAAACGGGA
CENP-B box, monomer 5   CTATGGTAGTAAAGGGA.

A 5'-biotinylated oligonucleotide was generated that matched the sequence of Bluescript KS+ upstream of the SmaI site: (5'-biotin-biotin-TCGAATTCCTGCAGCCC-3'). This oligonucleotide was used along with the T7 primer (Stratagene) in a polymerase chain reaction to generate double-stranded 5' biotinylated α-satellite DNA monomers that either contained or lacked the CENP-B box consensus sequence.

Preparation of COS Cell Extracts and Binding Reactions

COS-7 cells grown on 10-cm plates in DME with 10% FBS were transfected with 15 μg DNA using TransfectACE according to the manufacturer's instructions (Bethesda Research Laboratories, Gaithersburg, MD). Cells were exposed to the DNA for 12–24 h and then incubated in fresh complete media for up to an additional 24 h.

Transfected cells were washed twice with PBS, trypsinized, suspended in 5 ml media, concentrated by a brief low speed centrifugation, resus-

pended in 2 ml 75 mM KCl, and incubated 20 min on ice. The cells were then centrifuged again and resuspended in 1 ml lysis buffer (10 mM Tris-HCl, pH 8.3, 25 mM NaCl, 5 mM EDTA). Lysis was carried out by passing the cells through a 20-gauge needle with a bent tip 10–12 times, and was checked by phase contrast microscopy. One-fifth volume of 1.2 M NaCl was added to the lysate to bring the final concentration of NaCl to 200 mM. The lysate was then incubated on ice for 20 min, followed by a 30-min centrifugation in a microfuge at high speed. The supernatant was collected and used for binding reactions. Unused portions were quick-frozen in liquid nitrogen and stored at −70°C.

15 μl of the lysate was combined with an equal volume of 2× binding buffer (1× is 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0, 100 μg/ml BSA, 0.1% Triton X-100). The 2× buffer contained 0, 0.5, or 5 mg of salmon sperm DNA and 2 μl (∼3% of the PCR reaction) biotinylated substrate DNA amplified from p17M15 or p17M5. The final NaCl concentration in the binding reaction was 150 mM. Binding was carried out at room temperature for 5 min. 5 μl of Streptavidin agarose beads (Pharmacia Fine Chemicals, Piscataway, NJ; pre-equilibrated in 1× binding buffer containing 150 mM NaCl) was added to the binding reaction and incubated an additional 20 min at room temperature. The binding reactions were then centrifuged in a microfuge for 2 min. The supernatant was removed and mixed with 1/2 vol 3× SDS sample buffer (Earnshaw et al., 1985), boiled, and subjected to SDS-PAGE. The pellets were resuspended in 1× SDS sample buffer and boiled before SDS-PAGE.

Immunofluorescence

Immunofluorescence was performed essentially as described in Bernat et al. (1990) except that cells were fixed either with 3% formaldehyde or with cold methanol for 5 min. Antibody to the E1 tag was used at 1:1,000 dilution, and was detected with biotinylated anti–rabbit IgG (1:500 dilution) followed by streptavidin:Texas red (1:800 dilution). In double label experiments, human serum GS was used at 1:5,000 dilution, followed by FITC-conjugated anti–human IgG (1:200) to visualize centromeres. Chromosome spreads of transfected cells were prepared as described in Earnshaw et al. (1989). Hybridoma supernatant containing anti-TRPE mAbs, the generous gift of J. Harris (University of Florida, Gainesville, FL) was used at 1:10 dilution for immunofluorescence.

α-E1 Peptide Antibodies

A peptide encoding the carboxy-terminal 22 amino acids of the avian coronavirus E1 glycoprotein (gift of C. Machamer) was coupled to keyhole limpet hemocyanin with glutaraldehyde as described by Kreis (1986) and used as immunogen by Hazelton Research Products (Denver, PA) to generate rabbit polyclonal antibodies specific for the E1 tag. Initial experiments used affinity-purified rabbit antibodies generated to the same peptide coupled to BSA (gift of C. Machamer).

Protein Isolation, Gels, and Immunoblots

Populations of transfected cells were trypsinized, collected by a brief low speed centrifugation in a microcentrifuge, and total cellular protein was isolated from the resulting cell pellet by sonication in 200 μl SB followed by boiling for 5 min. SB is 50 mM Tris-HCl, pH 6.8, 15% sucrose, 2 mM EDTA, 3% SDS, 0.01% bromophenol blue, 20 mM DTT, 2.5 mM EGTA, pH 8.0, 350 μg/ml TAME, 0.5 mM PMSF, and 50 KIU/ml aprotinin. Electrophoresis and immunoblotting were as described in Earnshaw et al. (1984).

Results

Tagged Full-Length CENP-B Expressed in HeLa Cells Binds to Centromeres

To express CENP-B in human cells by DNA transfection, an epitope-tagged derivative of CENP-B was constructed in the mammalian expression vector pECE. This vector, which requires only an in-frame ATG start codon for expression (Ellis et al., 1986), contains the SV-40 origin, early and late promoters and polyadenylation signal, and a polylinker region followed by stop codons in all three reading frames. An immunological tag incorporated into this vector encodes the
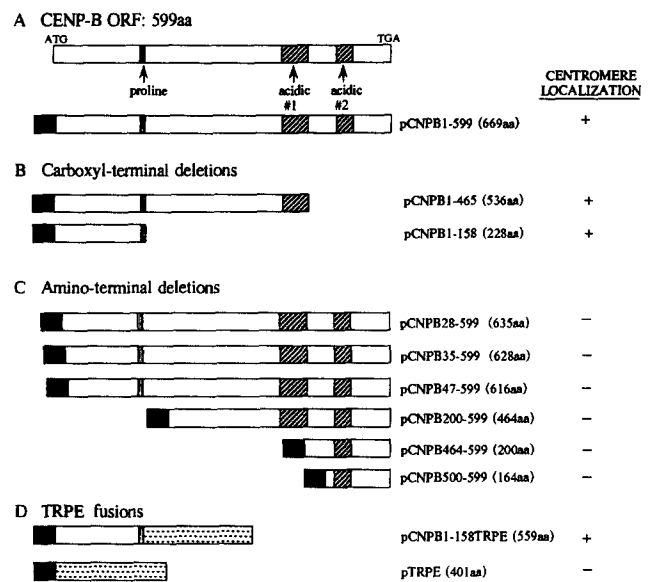

Figure 1. Schematic representation of CENP-B constructs and summary of immunofluorescence results. Localization of protein expressed by the various constructs was detected using anti-tag antibody, which specifically recognizes epitope(s) encoded by the E1 tag designated by the solid black box on each construct. + indicates centromere localization as verified by double staining with human anticentromere antibodies; − indicates that protein expressed by a construct does not localize to centromeres. The name of each construct indicates, in numbers, the amino acid residues of the CENP-B protein expressed. The total number of amino acid residues is indicated in parentheses. The regions of the TRPE fusion proteins encoding anthranilate synthetase in D are indicated by the stippled boxes.

carboxy-terminal 25 amino acids of E1, an avian coronavirus glycoprotein associated with membranes in infected cells (Machamer and Rose, 1987). This region of the E1 protein is not involved in either glycosylation or membrane binding. The E1 tag was engineered to contain its own ATG start codon and thus can initiate translation of any desired in-frame cDNA when present at the 5′ end of an open reading frame. Polyclonal antibodies made to a peptide derived from this tag sequence display little or no background staining of untransfected HeLa cells as determined by indirect immunofluorescence, thus allowing us to distinguish protein encoded by expression constructs from endogenous CENP-B in transfected human cells (e.g., see Fig. 4).

The efficacy of using such chimeric CENP-B proteins in these studies was initially tested by constructing plasmid pCNPB1-599, which contains the entire 599 amino acid open reading frame of CENP-B preceded by nine amino acids encoded by genomic upstream (5′) sequence and two direct repeats of the E1 tag, encoding an additional 56 amino acids at the extreme amino terminus (Fig. 1 A; see Materials and Methods). (We have found that detection of tagged CENP-B molecules in immunofluorescence experiments is facilitated by having multiple copies of the E1 sequence.)

When pCNPB1-599 was introduced into HeLa cells by transient DNA transfection, the expressed protein displayed a punctate nuclear staining pattern in interphase cells and a centromeric staining pattern in mitotic cells (Fig. 2). The expressed protein visualized with anti-tag antibody colocalized
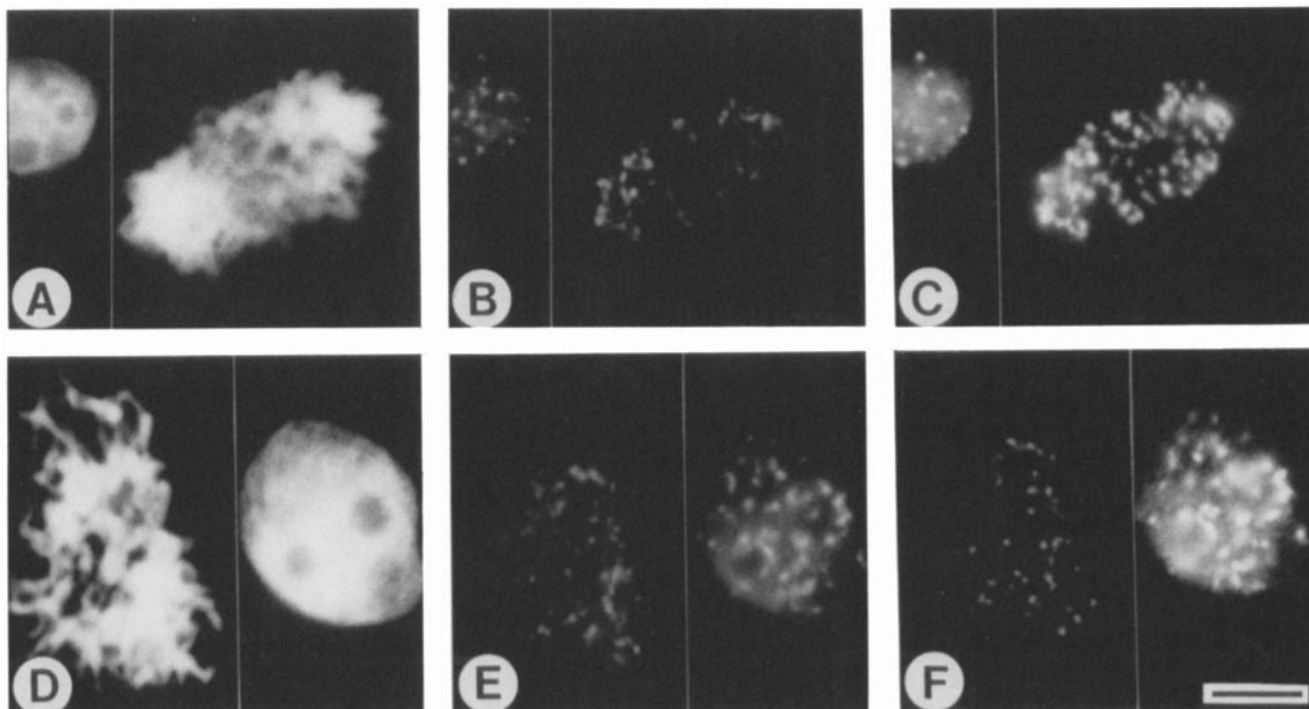
Figure 2. Tagged full-length and carboxyl-terminal deletion proteins localize to centromeres. HeLa cells transfected with pCNPBl-599 (*A–C*) or pCNPBl-158 (*D–F*) were processed for double immunofluorescence to visualize centromeres (*B* and *E*) with human anticentromere antibodies and the tagged CENP-B proteins with anti-tag antibody (*C* and *F*). *A* and *D* show staining of the DNA by DAPI. Examples of interphase (cell on left, *A–C*; cell on right, *D–F*) and mitotic (cell on right, *A–C*; cell on left, *D–F*) staining patterns are shown for each construct. Bar, 10 μm.

both with centromeres (defined using human anticentromere antibodies [Fig. 2, *A–C*]), and with endogenous CENP-B (defined using monoclonal CENP-B antibodies) (data not shown). Therefore, the E1 tag fusion does not appear to interfere with the stability, expression, or localization of CENP-B.

Expression of Carboxyl-Terminal Deletions Identifies a Centromere Localization Domain

Deletion derivatives of the tagged full-length CENP-B were constructed to determine if expression of such mutant proteins might disrupt or affect centromere function and/or assembly. These experiments revealed that the first 158 amino acids of CENP-B contain a signal responsible for localizing the protein to the centromere.

As a first step in the analysis, two carboxyl-terminal deletions of pCNPBl-599 were constructed: pCNPBl-465, which is deleted for all amino acids distal to the end of the first acidic domain, and pCNPBl-158, which is deleted for all amino acids distal to the proline motif (Fig. 1 *B*; also see Fig. 9). When either of these deletion constructs was transfected into HeLa cells and expression of chimeric protein was monitored by indirect immunofluorescence using anti-tag antibody, the signal detected in interphase cells was nuclear and punctate and colocalized with signal obtained with antibodies against centromeres. Remarkably, protein expressed by pCNPBl-158, the shortest tagged construct, in which the carboxyl-terminal three-fourths of CENP-B has been deleted, as well as that expressed by pCNPBl-465 (data not shown), localized to centromeres in mitotic cells with a pat-

tern that was indistinguishable from that produced by either tagged full-length CENP-B or by the endogenous CENP-B (Fig. 2, *D–F*).

It was readily apparent in the immunofluorescence experiments that the level of expression of exogenous (transfected) CENP-B within individual cells varied considerably, with some cells expressing higher than endogenous levels of the protein. When analyzed by immunoblotting of whole cell lysates, however, expression of the tagged CENP-B proteins appeared to be roughly comparable to that of endogenous CENP-B. We have examined the relative levels of expression between populations of cells transfected with different CENP-B constructs by immunoblot analysis of total cellular protein. Endogenous CENP-B, which is also expressed in these cells, can be used as an internal standard against which to estimate the amount of expression of each chimeric protein when blots of protein extracts are probed with a polyclonal antibody made against a bacterially expressed CENP-B fusion protein (Cooke et al., 1990).

Because CENP-B migrates anomalously in SDS-PAGE (Earnshaw et al., 1987), it is not possible to make exact predictions for the gel mobilities of subcloned portions of the molecule. Despite this fact, we have noted that all E1-tagged constructs express proteins at least as large as would be predicted from the DNA sequence (Fig. 3). We thus conclude that all constructs (which were verified by restriction mapping) are transcribed and translated as expected. Multiple forms of tagged protein observed on immunoblots probably reflect alternative translational starts facilitated by the closely spaced ATGs in the two tags.
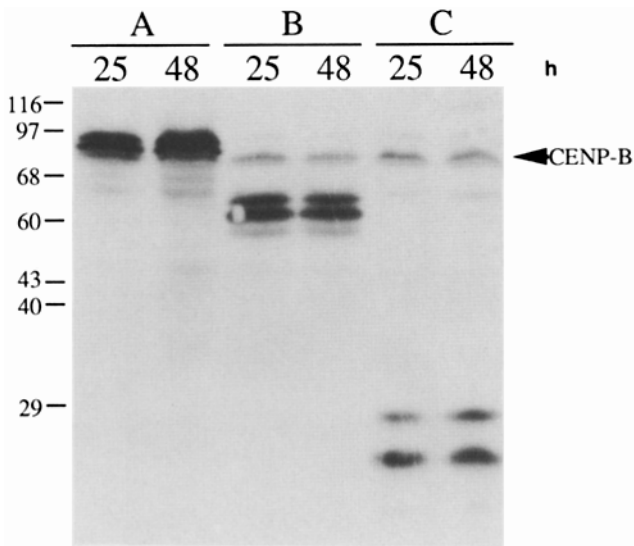
Figure 3. Immunoblot analysis of total protein isolated from transfected cells. Total cellular protein was isolated after 25 and 48 h of expression from cells transfected with pCNPB1-599 (*A*), pCNPB1-465 (*B*), and pCNPB1-158 (*C*). Tagged CENP-B proteins expressed by these constructs were detected after blotting to nitrocellulose with a polyclonal antibody made to a bacterially expressed CENP-B fusion protein. The band corresponding to endogenous CENP-B is indicated.

Fig. 3 also demonstrates the relative levels of expression of proteins encoded by the various transfected constructs. Densitometry of the lanes representing 25 h of expression revealed that protein expressed by pCNPB1-599, pCNPB1-465, and pCNPB1-158 are at least several-fold more abundant than endogenous CENP-B. However, since not every cell in the population is transfected, the ratio of the transfected signal to the endogenous signal represents an underestimate of expression in any given transfected cell relative to the expression of endogenous CENP-B.

In an alternative approach, we have fused the amino-terminal 131 residues of CENP-B in frame with residues 367–599. This protein is expressed normally and localizes to centromeres in both interphase and mitotic cells (data not shown). We will demonstrate below that no further centromere localization signals exist outside of domain I. This experiment thus serves to further delineate the centromere localization signal as being located within the amino-terminal 131 residues of CENP-B.

The CENP-B Centromere Localization Domain Directs a Noncentromere Protein to Centromeres

All of the tagged carboxyl-terminal deletion constructs whose expression was readily detectable on immunoblots also produced a centromere staining pattern by immunofluorescence. It was therefore necessary to eliminate the possibility that the tag itself might be responsible for the centromere localization of chimeric proteins.

The following experiments demonstrated that the E1 tag alone is unable to direct a noncentromere protein to centromeres. An expression construct was made in which the bacterial TRPE gene was fused in-frame downstream of DNA encoding the tandem E1 tags and the nine amino acids of CENP-B genomic upstream sequence present in all CENP-B

expression constructs. This construct, pTRPE (Fig. 1 *D*), was used to transfect HeLa cells and its expression was assayed by indirect immunofluorescence with anti-tag antibody. The fusion protein in these cells, although predominantly nuclear, was distributed diffusely and was not specifically localized to centromeres, chromosomes, or any other nuclear substructure (Fig. 4, *J–L*). Similar results were seen regardless of whether the cells were fixed with cold methanol or formaldehyde, although additional cytoplasmic staining was seen in the methanol-fixed cells. While the mechanism for this nuclear localization of the TRPE fusion protein is unclear, the protein is small enough ($\sim$44 kD; see Fig. 6 *A*) to undergo passive diffusion through nuclear pores (Bonner, 1978). These results confirmed that although an E1-tagged TRPE fusion protein was able to gain entry to the nucleus, the E1 tag could not localize the protein to centromeres. Thus it is highly unlikely that the tag contributes to localization of recombinant CENP-B variants to centromeres.

A similar approach was used to demonstrate that the centromere-localization domain of CENP-B identified above (by pCNPB1-158) is sufficient to direct a noncentromere protein to centromeres. To this end, the bacterial TRPE gene was fused in-frame downstream of the coding sequence in pCNPB1-158, and resulting construct, pCNPB1-158TRPE (Fig. 1 *D*) was introduced into HeLa cells by transfection. Immunoblot analysis of total protein isolated from transfected cells confirmed that a protein consistent in size with the expected molecular weight for the fusion protein was expressed in the transfected cell population (see Fig. 6 *A*). Indirect immunofluorescence using anti-tag antibody to detect the protein expressed by pCNPB1-158TRPE yielded a staining pattern consistent with centromere localization in interphase cells regardless of the fixation procedure used (Fig. 4, *A–F*). Furthermore, centromere staining was detected even when the transfected cells were stained with a mAb that recognized the TRPE portion of the CENP-B fusion protein. This latter pattern also colocalized with signals obtained with human anticentromere antibodies, indicating that the apparent centromere localization of pCNPB1-158TRPE is a property of the complete fusion protein and is not due to a processed subset of the protein population lacking the TRPE moiety (Fig. 4, *G–I*).

The centromeric localization of the pCNPB1-158TRPE protein was seen particularly clearly when chromosome spreads were prepared from transfected cells arrested in mitosis. Because few mitotic cells expressing protein encoded by this construct were observed in the original experiments, transfected cells were blocked in mitosis with the microtubule depolymerizing drug colcemid, and spreads of mitotic chromosomes were prepared by methanol/acetic acid fixation for indirect immunofluorescence using antitag antibody. Fig. 5 clearly demonstrates that the CENP-B/TRPE fusion protein expressed by pCNPB1-158TRPE localizes specifically to centromere regions of mitotic chromosomes.

These results demonstrate that the centromere localization signal on CENP-B can be used to redirect noncentromere proteins to centromeres, similar to the methodology previously exploited for analysis of nuclear localization signals (Kalderon et al., 1984; Robbins et al., 1991).

CENP-B Contains a Unique Centromere Localization Domain

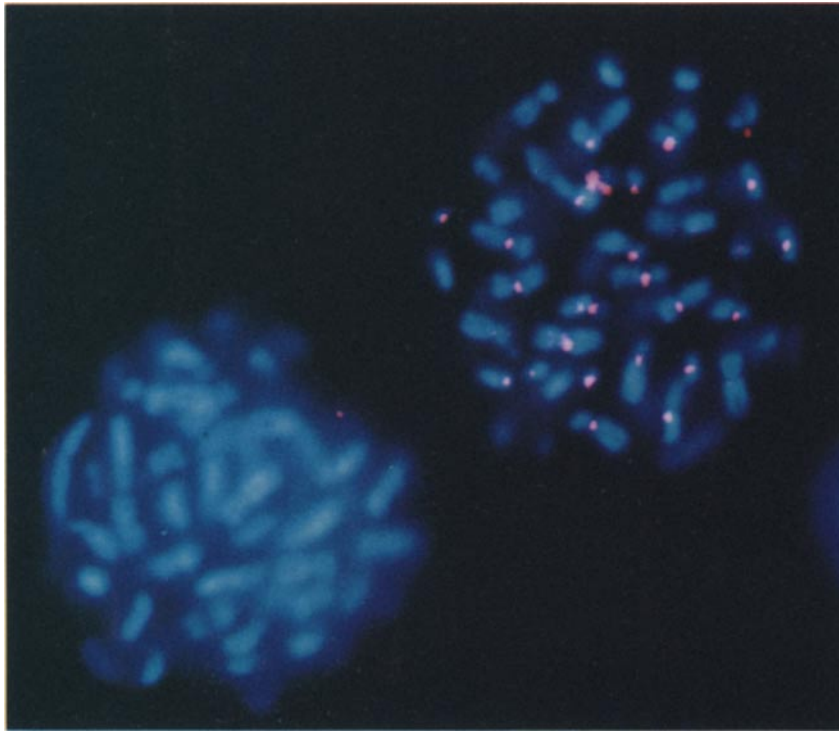Although results from the previous set of experiments ruled
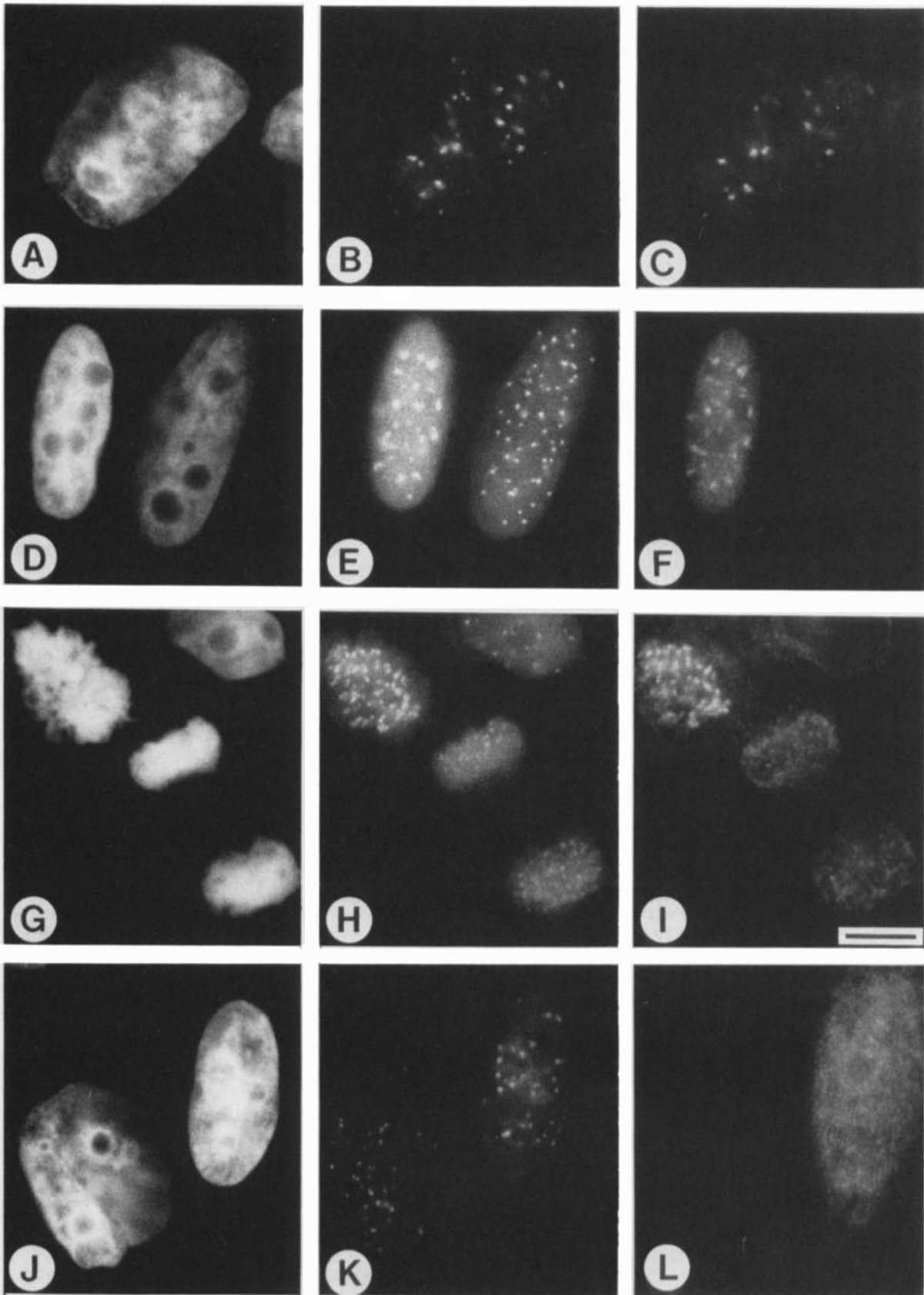
Figure 5. Chromosome spreads showing the ability of the amino-terminal domain of CENP-B to localize an exogenous protein to centromeres. Cells transfected with pCNPB1-158TRPE were spread after methanol/acetic acid fixation and stained with the antibody to the E1 tag using a previously described procedure (Earnshaw et al., 1989). The variability of staining between different chromosomes is reminiscent of that observed when antibodies monospecific for CENP-B are used to detect the endogenous protein (Earnshaw et al., 1987). The cell at the lower left was not transfected.

out involvement of the E1 tag in centromere localization, they did not eliminate the possibility that multiple centromere localization signals are present on CENP-B. To clarify this point, and to determine what effect, if any, expression of carboxyl-terminal portions of CENP-B might have on centromere function and/or assembly, a series of in-frame amino-terminal deletions of CENP-B was constructed and introduced into HeLa cells by transfection (Fig. 1 C). All deletion constructs were tagged at the amino terminus with the tandem E1 tag and examined by indirect immunofluorescence with antitag antibody (see Fig. 1 C and Materials and Methods).

The only region of CENP-B capable of localizing the protein to the centromere was that contained in the first 158 amino acids as defined by pCNPB1-158. Deletions that removed 28, 35, 47, 200, 464, and 500 amino acids from the amino terminus of CENP-B abolished the ability of the protein to assemble at chromosomal centromeres but did not prevent these proteins from accumulating in the nucleus (data not shown). All of these constructs produced proteins consistent with the expected sizes when total protein isolated from transfected cells was subjected to immunoblot analysis (Fig. 6 B).

We have examined the distribution of protein encoded by the smallest amino-terminal deletion (pCNPB28-599; Fig. 1 C) in the greatest detail. This construct deletes ∼5% of

the CENP-B open reading frame, which corresponds to ∼18% of the centromere localization domain identified by pCNPB1-158. Protein expressed by pCNPB28-599 displayed a diffuse staining pattern in interphase cells that was predominantly nuclear although some slight cytoplasmic staining was also observed (data not shown). Numerous examples of mitotic cells expressing high levels of the truncated protein were also observed. In no case, however, did this protein localize to centromeres of mitotic chromosomes (confirmed by double staining with human anticentromere antibodies). Instead the protein displayed a diffuse staining throughout the cytoplasm of these cells (Fig. 7).

These experiments demonstrate that the amino-terminal 28 amino acids of CENP-B are essential (though possibly not sufficient) for the localization of this protein to centromeres. This rules out the possibility that other centromere localization signals exist outside of the first 131 amino acids on CENP-B. No additional phenotypes resulting from the expression of any carboxyl-terminal portions of CENP-B were observed in these experiments.

Centromere Localization Domain Displays "CENP-B Box"-specific DNA Binding Activity

We wished to further characterize the centromere localization signal located within the amino-terminal 158 amino

Figure 4. The amino-terminal domain of CENP-B can localize bacterial anthranilate synthetase to centromeres. HeLa cells transfected with pCNPB1-158TRPE (*A-I*) or pTRPE (*J-L*) were processed for double immunofluorescence to visualize centromeres and the E1-tagged proteins. *A*, *D*, *G*, and *J* show staining of the DNA by DAPI. *B*, *E*, *H*, and *K* show the centromere staining pattern obtained using human anticentromere antibodies. *C*, *F*, *I*, and *L* show the distribution of protein expressed by the transfected constructs as detected with anti-tag antibody (*C*, *F*, and *L*) or anti-TRPE mAb (*I*). Cells in *A-C* and *J-L* were fixed with cold methanol, while those shown in *D-I* were fixed with 3% formaldehyde. The cell on the right in *F* does not stain with anti-tag antibody presumably because it was not transfected with the chimeric construct. Bar, 10 μm.
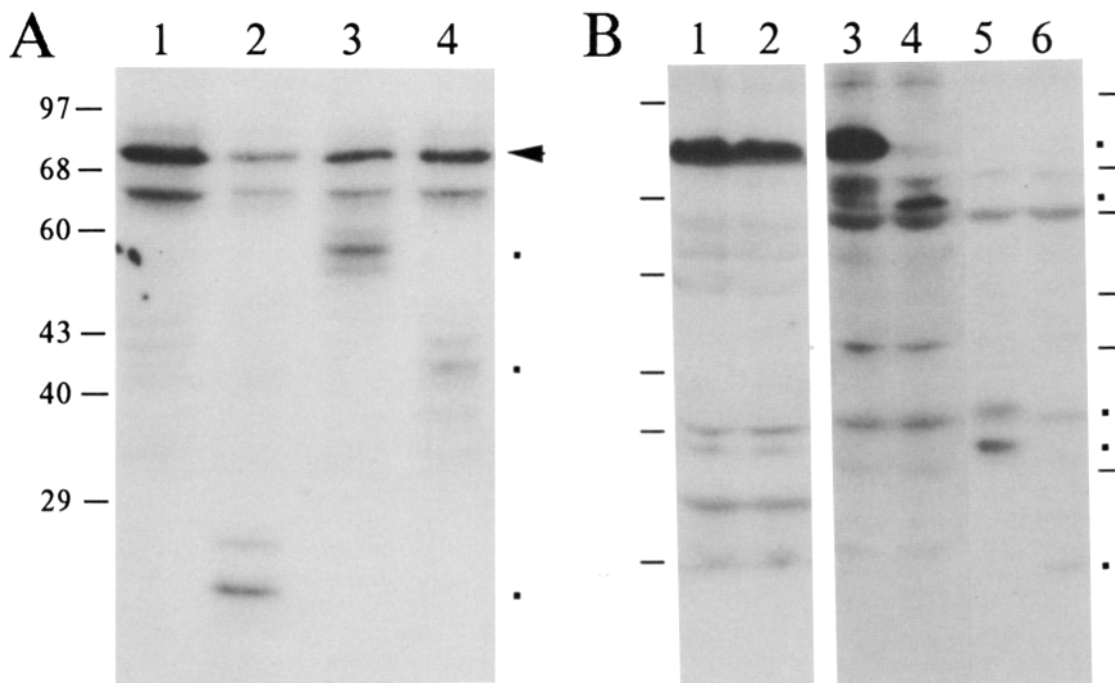
Figure 6. Immunoblot analysis of total protein isolated from cells transfected with TRPE fusion constructs (*A*) and amino terminal deletion constructs (*B*). The protein blot in *A* was probed with a polyclonal antibody made to a CENP-B fusion protein produced in a pATH vector. Thus, this antibody detects epitopes for CENP-B as well as those for TRPE. Lane *1*, untransfected cells; lane *2*, cells transfected with pCNPB1-158; lane *3*, cells transfected with pCNPB1-158TRPE; and lane *4*, cells transfected with pTRPE. The positions of novel bands corresponding to the expressed proteins are indicated by the black dots to the right of each panel. The band corresponding to endogenous CENP-B present in all lanes is indicated by the arrowhead. The protein blot in *B* was probed with the anti-tag antibody. Protein was isolated from cells transfected with the following constructs: lane *1*, pCNPB28-599; lane *2*, pCNPB35-599; lane *3*, pCNPB47-599; lane *4*, pCNPB200-599; lane *5*, pCNPB464-599; and lane *6*, pCNPB500-599. The positions of novel bands corresponding to the expressed proteins are indicated by the black dots. The positions of molecular weight markers having the same relative sizes shown in *A* are indicated by the short horizontal lines; phosphorylase b (97 kD), BSA (68 kD), catalase (60 kD), actin (43 kD), aldolase (40 kD), and carbonic anhydrase (29 kD).

acids of CENP-B. The following experiments were designed to ask if this region of CENP-B is able to interact specifically in vitro with α-satellite DNA monomers containing a consensus CENP-B box sequence (Masumoto et al., 1989).

To determine the binding properties of truncated versions of CENP-B, we needed an experimental system that could generate abundant quantities of the mutant proteins in soluble form for in vitro DNA binding assays. We chose to use African green monkey (COS) cells as the recipient (host) for transfection because they offered four significant advantages for these experiments. First, they do not appear to express any form of endogenous CENP-B, as judged by immunofluorescence, using any of a number of polyclonal antibodies (W. C. Earnshaw, unpublished; J. B. Rattner, personal communication). Therefore, transfection of COS cells with mutant human CENP-B would not result in competition between endogenous and exogenous protein for specific binding to the DNA substrate. Second, the transfected CENP-B is expected to remain soluble in COS cells, since African green monkey satellite DNA lacks the CENP-B box (J. B. Rattner, personal communication). Third, COS cells constitutively express large T antigen, which enables replication of SV-40 origin-containing plasmids, making it possible to express particularly high levels of protein from cDNAs cloned into pECE vectors. Fourth, any posttranslational modifications associated with eukaryotes will be made in this expres-

sion system. Thus, COS cells provide a unique eukaryotic system for generating extracts for binding experiments that lack endogenous CENP-B and contain mutant versions of soluble CENP-B that are not bound to α-satellite DNA.

Whole cell extracts were prepared from COS cells transfected with either pCNPB1-158 or pCNPB47-599 (Fig. 1), and the expressed truncated CENP-B proteins were solubilized and mixed with the DNA substrates. The substrates for the DNA binding reactions were amplification products resulting from polymerase chain reaction of individually cloned α-satellite DNA monomers that either contain or lack the 17-bp CENP-B box consensus sequence (Masumoto et al., 1989; Waye and Willard, 1986). The amplified α-satellite monomer DNAs were labeled at one end with biotin, which was covalently linked to one of the PCR primers. The biotin label allowed the use of streptavidin agarose to precipitate the DNA, along with any associated protein, from the binding reactions. Subsequent fractionation of the expressed protein was detected by immunoblot analysis using antitag antibody. Protein capable of specific binding to the DNA was present in the pellet fraction. Protein lacking DNA binding activity was present in the supernatant.

Fig. 8 shows the results obtained from DNA binding experiments using extracts made from cells expressing pCNPB1-158 (panels *B*). In the absence of any added α-satellite monomer or competitor DNA, the expressed protein remains in
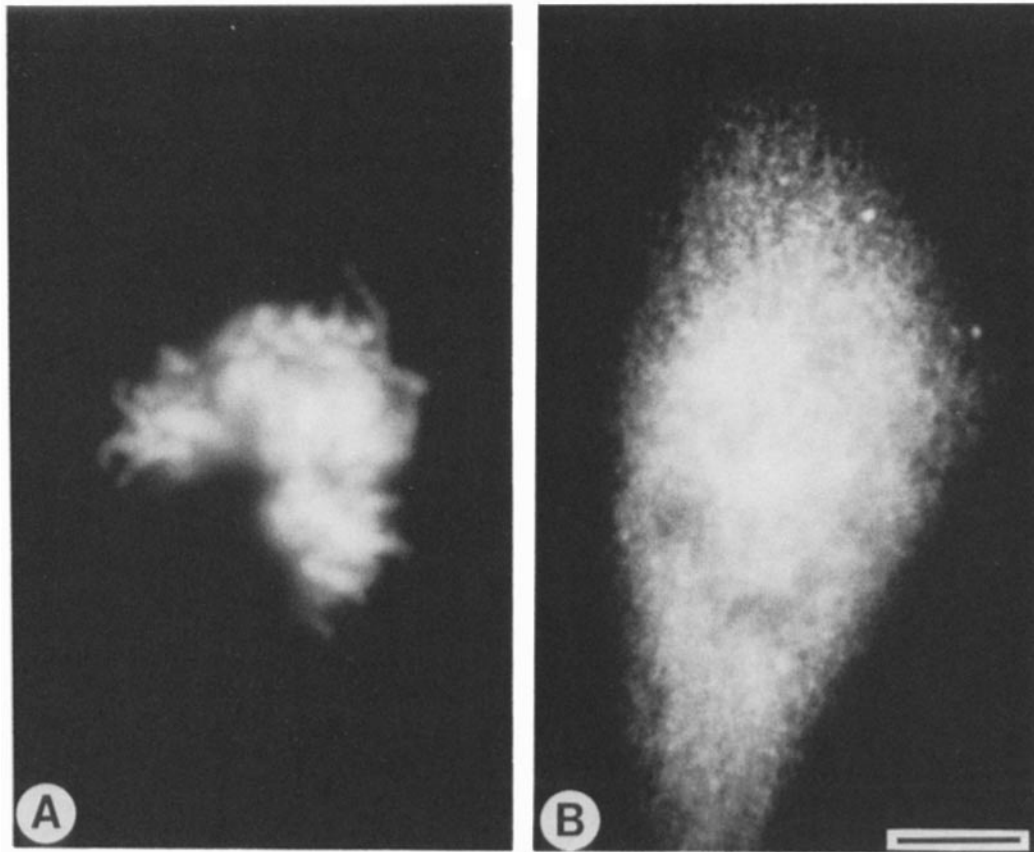
Figure 7. Removal of 28 amino acids from the amino terminus of CENP-B abolishes centromere localization. A mitotic HeLa cell transfected with pCNPB28-599. (*A*) DNA visualized by DAPI staining. (*B*) Protein expressed by pCNPB28-599 visualized with anti-tag antibody. Bar, 10 μm.

the supernatant (unbound) fraction (panel *B*, lane *4*). The same result was obtained when protein expressed by this construct was incubated with a biotinylated α-satellite DNA monomer lacking a consensus CENP-B box sequence (panel *B*, lanes *11-16*). However, when protein expressed by pCNPB1-158 was incubated with a biotinylated α-satellite DNA monomer containing a consensus CENP-B box sequence, the protein was detected in the pellet fraction (panel *B*, lanes *5-10*). The absence of significant signal in the supernatant upon addition of increasing amounts of nonspecific competitor DNA is indicative of the high affinity of this protein for DNA containing the CENP-B box.

The preceding experiment demonstrated that a polypeptide consisting of only the amino-terminal 158 amino acids of CENP-B could bind specifically to DNA containing the CENP-B box. This binding was abolished if the amino-terminal 47 amino acids were removed from CENP-B. In these experiments, protein expressed by pCNPB47-599 remained in the supernatant (unbound) under all conditions examined (panels *A*, lanes *1-16*). As previously discussed, protein encoded by pCNPB47-599 failed to localize to centromeres in transfected HeLa cells (data not shown).

Therefore, the ability of CENP-B to localize to centromeres in vivo correlates with the ability of the protein to bind a centromere-specific DNA sequence in vitro. These experiments suggest that CENP-B uses a direct mechanism, mediated by DNA binding to the CENP-B box, for localization to the centromere.

Discussion

CENP-B Is a Multidomain Protein

The deduced amino acid sequence of CENP-B (Earnshaw et al., 1987) suggests that this protein is comprised of four domains separated by three hinge regions (Fig. 9). The most prominent of these putative hinge regions, A1 and A2, are characterized by extremely high concentrations of glutamic acid (glu) and aspartic acid (asp) residues, making it unlikely that either assumes a single stable secondary structure. Therefore, we have suggested that both A1 and A2 are probably unstructured hinge regions (Pluta et al., 1990). An additional putative hinge region separates domains I and II. This proline-rich motif (P-X-X-P)$_4$ also occurs in MAP2 and tau (Joly et al., 1989), where it has been shown to be a site of enhanced protease sensitivity. This is not the microtubule-binding region of either MAP2 or tau (Joly et al., 1989). Domain I of CENP-B is absolutely conserved between the human and mouse homologues of this protein, and is distinguished by a basic DNA-binding motif characteristic of the helix–loop–helix protein family (Sullivan and Glass, 1991).

In this report, we describe the first detailed structure/function analysis of CENP-B, a protein of the human centromeric heterochromatin. Three previous results provide a framework within which to think about the role of CENP-B in centromere structure and kinetochore organization. Firstly, in vitro CENP-B binds specifically to a subset of α-satellite
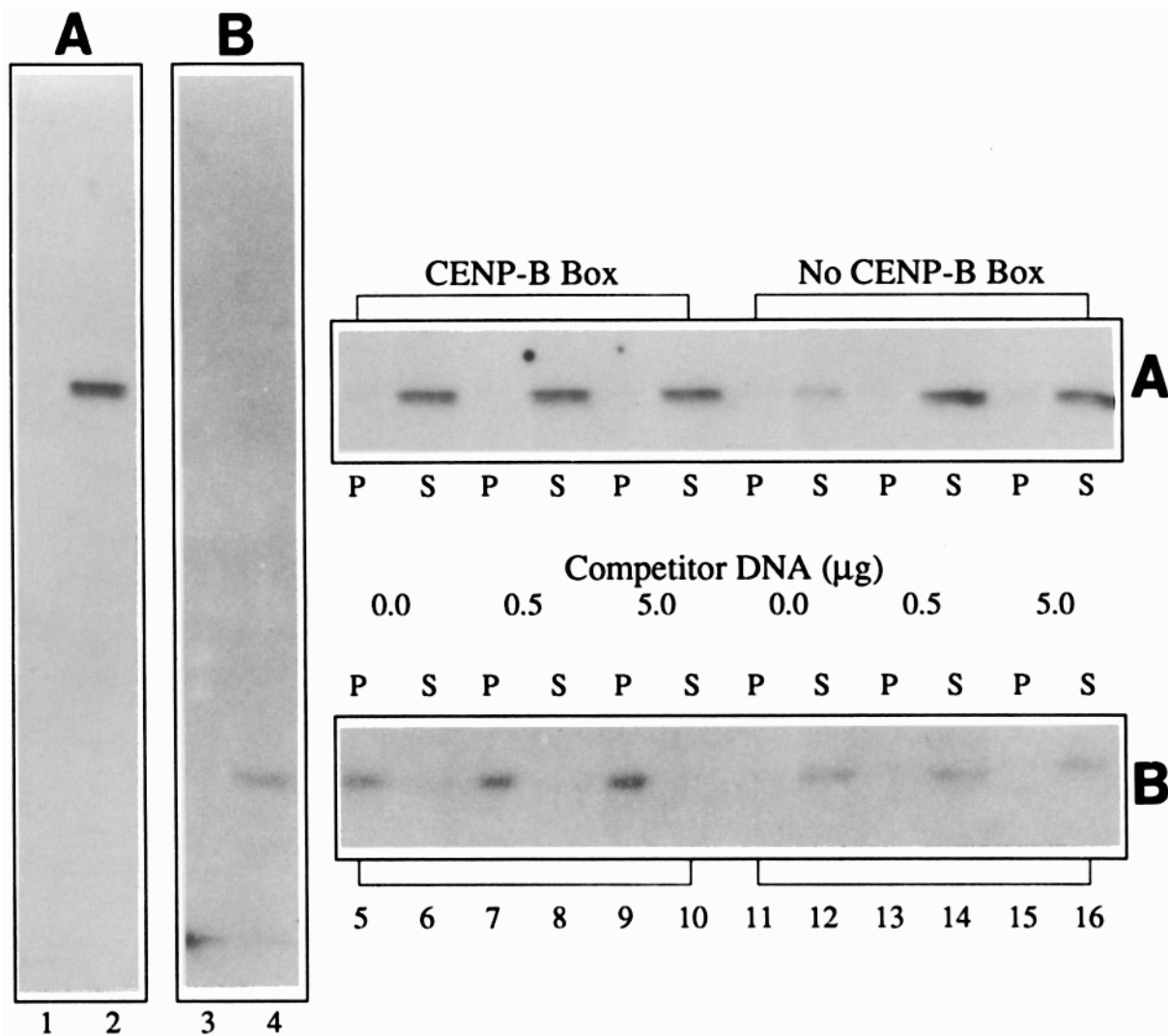
Figure 8. The amino-terminal 158 amino acids of CENP-B bind specifically to individual human α-satellite DNA monomers containing the CENP-B box in vitro. The CENP-B proteins used for the binding were obtained from extracts of African green monkey cells transfected with the truncated CENP-B proteins pCNPB1-158 or pCNPB47-599 (Fig. 1). Individual biotin-labeled α-satellite DNA monomers either containing or lacking the CENP-B box were incubated with extracts expressing the truncated CENP-B proteins pCNPB47-599 (panels *A*) or pCNPB1-158 (panels *B*). Streptavidin-agarose beads were introduced into the reactions to precipitate the DNA and associated proteins. The resulting pellets (*P*) and supernatants (*S*) were subjected to SDS-PAGE followed by immunoblotting using antibody against the E1 tag. Analysis of the pellets and supernatants from control reactions performed in the absence of DNA (lanes *1-4*) show that the proteins do not adhere to the agarose beads; lanes *1* and *3* are pellets, lanes *2* and *4* are supernatants. The anti-E1 antibody recognizes a single species in each extract. Lanes *5-10* are reactions with p17M15 (α-satellite DNA monomer containing the CENP-B box), and lanes *11-16* are with p17M5 (α-satellite DNA monomer lacking the CENP-B box). Competitor DNA (sonicated herring sperm DNA) was used in amounts ranging from 0 to 5.0 μg in a 30-μl reaction.

DNA monomers carrying the 17-bp CENP-B box (Masumoto et al., 1989). Secondly, in human chromosomes both CENP-B and the other centromeric autoantigens are located in the heterochromatin beneath the outer kinetochore plate (Cooke et al., 1990). Thirdly, antibody microinjection experiments suggest that the CENP antigens are essential for the assembly of a functional kinetochore at mitosis (Bernat et al., 1991).

Specific DNA Binding by Domain I Directs CENP-B to the Centromere

CENP-B has attracted interest primarily because of its specific localization at the centromere region of human chromosomes. As will be discussed at greater length below, the

physiological role of the protein is unknown. However, a number of considerations suggest that CENP-B is likely to be involved in heterochromatin structure, where it may influence formation of the primary constriction, assembly of the kinetochore, or even pairing of sister chromatids.

In principle, two sorts of interactions could contribute to the assembly of a protein such as CENP-B onto the centromere. The most straightforward of these is via the recognition of specific DNA sequences. Alternatively, protein–protein as well as protein–DNA interactions may be required. The latter possibility is consistent with our earlier demonstration that CENP-B is a component of the mitotic chromosome scaffold fraction (Earnshaw et al., 1984). Components
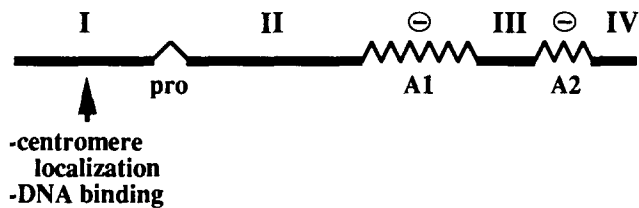
Figure 9. Putative domain organization of CENP-B.

of the chromosome scaffold are widely thought to have a structural role in mitotic chromosomes (Laemmli et al., 1978).

Our present experiments argue strongly that CENP-B localization at the centromere is directed by the specific recognition of the CENP-B box DNA by domain I of the protein. We have clearly shown that domain I is both necessary and sufficient for localization of CENP-B to centromeres in vivo and for specific recognition of DNA segments containing the CENP-B box in vitro. Furthermore, deletion of only 28 amino acids from domain I, which removes much of helix 1 of the helix–loop–helix motif (Sullivan and Glass, 1991), renders CENP-B unable to bind to the centromere although it is still able to enter the nucleus. This indicates that other interactions (if any) made by domains II–IV of CENP-B are not sufficient to localize the protein to the centromere.

CENP-B May Be Functionally Redundant

Our present results suggest that the localization of CENP-B to the centromere is determined solely by the binding of the protein to CENP-B boxes in the α-satellite DNA. Thus the distribution of CENP-B in the heterochromatin can be modeled based on what is known about the distribution of CENP-B boxes. Examples of α-satellite monomers from all of the human chromosomes have been cloned and sequenced. These monomers (consensus length of 171 bp) exist in tandem arrays that form higher-order repeats, called alphoid subfamilies. Recently, the sequences of 293 different α-satellite monomers, representing 33 different subfamilies from 24 human chromosomes, have been compiled into an overall consensus sequence (Choo et al., 1991). Analysis of the statistical data shows that within each monomer there is only one position where there is a significant probability of encountering a CENP-B box-position 127 of the overall consensus. Given the actual frequency of nucleotides observed at each position of the consensus (Choo et al., 1991), the probability of a CENP-B box occurring in a given α-satellite monomer is 0.17%, or approximately once in 600 monomers. In fact, the CENP-B box was found to appear significantly more frequently than this in one selected group of cloned α-satellite monomers (Masumoto et al., 1989).

It is puzzling that both CENP-B and the CENP-B box have yet to be detected on the human Y chromosome (Earnshaw et al., 1991). The most reasonable solution to the paradox is that CENP-B is functionally redundant and that its role is shared with another chromosomal component. If CENP-B activity is indeed important, and centromeres of Y chromosomes function in the same way as do those of other chromosomes, then there must be a CENP-B homologue present on the Y chromosome that is undetectable with our polyclonal antibodies and that binds to a sequence distinct from the CENP-B box.

This redundancy may explain why we have been unable to observe a mitotic or cell cycle arrest phenotype in cells transfected with any of the constructs shown in Fig. 1. When cultures transiently expressing proteins encoded by these constructs were observed by immunofluorescence microscopy, transfected cells in all phases of mitosis were detected. This was surprising, since the most obvious model for CENP-B function is that binding of domain I to α-satellite DNA permits the (presumably) mobile domains II–IV to contact other components of the chromosome (Fig. 9). If true, this predicts (as suggested by Herskowitz, 1987) that occupancy of DNA binding sites by protein containing only domain I might be detrimental. In this case, DNA binding functions would be uncoupled from any functions normally carried out by the other CENP-B domains.

Expression of isolated CENP-B subdomains might have failed to lead to observable phenotypes in these experiments because the structural and functional redundancy of the centromeric heterochromatin may give the system a significant "buffering" capacity with regard to the expression of portions of CENP-B, so long as a certain level of endogenous CENP-B is present. Endogenous CENP-B is present in several hundred copies per centromere (Bernat, 1991). It is unlikely that the failure to observe a phenotype was because of insufficient expression of the transfected proteins. Expression was found to vary greatly across the cultures, with some cells showing a significant overproduction of the transfected proteins. Cells over-expressing CENP-B domains were observed in all phases of (apparently) normal mitosis.

The effect of expressing CENP-B domains could be complicated if the helix–loop–helix motif of domain I was involved in dimerization as well as DNA binding (Murre et al., 1989). In this case, mutant proteins might bind to DNA as heterodimers, (i.e., one endogenous wild-type polypeptide and one transfected mutant polypeptide), which may be partly or fully functional.

Order and Disorder in the Centromere: Possible Random-sampled Modification of Chromatin Structure by CENP-B

Our results suggest that the binding of CENP-B to centromeres in vivo is apparently determined solely by sequence-specific recognition of the CENP-B box. Thus, the distribution of the protein along the satellite DNA should reflect the frequency of occurrence of the box motif. Since alphoid arrays from different chromosomes have different patterns of occurrence of the CENP-B box motif (Masumoto et al., 1989), the distribution of CENP-B will also necessarily differ.

The distribution of CENP-B on α-satellite DNA may be modeled as the random occupancy (or sampling) of a one-dimensional lattice. In this case, the lattice constant is 171 bp, the length of the α-satellite monomer. Thus, although the probability of binding of CENP-B to any given monomer will be low (since most monomers lack CENP-B boxes), any CENP-B molecules that are bound must exist in locally identical environments. Furthermore, all CENP-B molecules will be spaced at intervals of $n \times 171$ bp, where n may be any integer, but will typically be between 4 and 35 (the range of alphoid array sizes observed to date; Choo et al., 1991). Such a distribution has interesting potential consequences for the chromatin structure of the centromere.
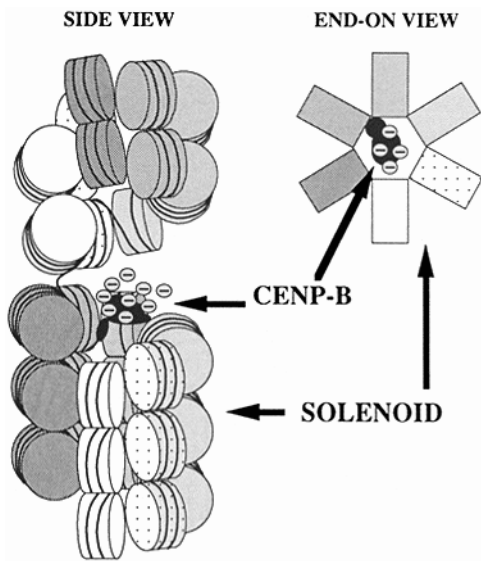
SIDE VIEW   END-ON VIEW

← CENP-B

← SOLENOID

Figure 10. Binding of CENP-B to α-satellite chromatin is likely to affect the structure of the 30-nm fiber (here modeled as a solenoid [Widom and Klug, 1985]). Assuming that the phasing of nucleosomes in human α-satellite chromatin resembles that in the African green monkey, the CENP-B box will be located on, or adjacent to, the internucleosomal linker. For nucleosomes with such a short linker (∼6 bp), the CENP-B box is thus likely to be found in the interior of the solenoid. Binding of a highly negatively charged protein such as CENP-B at this position may disrupt the continuity of the solenoid, as shown.

The nucleosomal structure of African green monkey α-satellite DNA has been studied extensively. Nuclease digestion experiments reveal that nucleosomes in α-satellite chromatin show a 172-bp periodicity, corresponding to the length of the consensus monomer repeat. Published evidence suggests that although several different nucleosome phasings occur in α-satellite chromatin, a significant fraction (∼35%) of the nucleosomes are phased with the inter-nucleosomal linker (as defined by nuclease sensitivity) centered around nucleotide 126 of the monomer consensus (Musich et al., 1982; Zhang et al., 1983). The overall degree of conservation of the α-satellite DNA sequence between human and monkey suggests that both species may maintain a constant nucleosome phase on this chromatin.

If human nucleosomes are phased on alphoid chromatin, then CENP-B will occupy a fixed spatial domain of the solenoidal fiber (although its occupancy along the length of the fiber will be random). The CENP-B box begins at position 127 of the human α-satellite monomer consensus. The nucleosome phasing results thus suggest that CENP-B will be localized immediately adjacent to, or within, the linker DNA between core particles. In this case, CENP-B would bind in the interior of the solenoid, particularly if one considers solenoidal models with an internally located linker (Fig. 10). This configuration is likely for α-satellite chromatin, where linker lengths are minimal.

Binding of CENP-B to the interior of the solenoid could affect the overall structure of the fiber. The two acidic domains of CENP-B together contain 76 glu and asp residues. It is entirely conceivable that presence of this concentration of negative charge in the interior of the solenoid might cause a significant disruption of the solenoidal lattice, as suggested

in Fig. 10. Thus CENP-B may function as a "solenoid-breaker". In regions where CENP-B boxes occur with a frequency of greater than one in six nucleosomes (one turn of the solenoid) (Finch and Klug, 1976), this could even preclude the formation of a classical solenoid structure. While this might seem to be the opposite of what would be expected for heterochromatin, it is possible that chromatin containing bound CENP-B uses some other mechanism of condensation (such as looping) involving interactions between nonhistone proteins.

Conclusions

Experiments reported here reveal that the localization of a major human heterochromatin protein, CENP-B, to the centromere is directed solely by a domain of <158 amino acids from the amino terminus of the molecule. This domain is responsible for the sequence-specific binding of the protein to α-satellite DNA monomers containing the CENP-B box motif in vitro. CENP-B is a component of the mitotic chromosome scaffold fraction, and we would have predicted that the assembly of this protein into the centromere might involve protein–protein contacts in addition to protein–DNA contacts. Surprisingly, however, our experiments are consistent with the primary mechanism for localization of this protein to the centromere being the specific recognition of a target DNA sequence, the CENP-B box.

One result of the arrangement of CENP-B and nucleosomes deduced from the analysis of the distribution of CENP-B boxes (Masumoto et al., 1989) is that the centromeric heterochromatin may possess an unusual degree of structural heterogeneity. Given the tremendous uniformity of DNA sequence in α-satellite DNA, centromeric heterochromatin has the potential to form a highly ordered solenoidal structure that could extend over millions of base pairs. At the same time, the CENP-B boxes may induce local perturbations of the solenoid lattice at intervals along its length that are directed by the underlying higher-order alphoid repeat (which differs from chromosome to chromosome). The net result may be that the structure of centromeric heterochromatin differs significantly from that predicted for bulk chromatin by the solenoid model.

References

Bernat, R. L. 1991. Human centromere proteins CENPs A, B and C: an in vivo analysis of centromere structure and function. Ph.D. thesis, Johns Hopkins University School of Medicine, Baltimore, MD. 128 pp.

Bernat, R. L., G. G. Borisy, N. F. Rothfield, and W. C. Earnshaw. 1990. Injection of anticentromere antibodies in interphase disrupts events required for chromosome movement at mitosis. *J. Cell Biol.* 111:1519–1533.

Bernat, R. L., M. R. Delannoy, N. F. Rothfield, and W. C. Earnshaw. 1991. Disruption of centromere assembly during interphase inhibits kinetochore

morphogenesis and function in mitosis. *Cell.* 66:1229-1238.

Bonner, W. 1978. Protein migration and accumulation in nuclei. *In* The Cell Nucleus. Chromatin, Part C. H. Busch, editor. Academic Press, Inc., New York. 97-148.

Brinkley, B. R., and E. Stubblefield. 1966. The fine structure of the kinetochore of a mammalian cell in vitro. *Chromosoma (Berl.).* 19:28-43.

Brown, S. W. 1966. Heterochromatin. *Science (Wash. DC).* 151:417-425.

Choo, K. H., B. Vissel, A. Nagy, E. Earle, and P. Kalitsis. 1991. A survey of the genomic distribution of alpha satellite DNA on all human chromosomes, and derivation of a new consensus sequence. *Nucleic Acids Res.* 19:1179-1182.

Cooke, C. A., R. L. Bernat, and W. C. Earnshaw. 1990. CENP-B: a major human centromere protein located beneath the kinetochore. *J. Cell Biol.* 110:1475-1488.

Daniels, S. B., M. McCarron, C. Love, S. H. Clark, and A. Chovnick. 1986. The underlying bases of gene expression differences in stable transformants of the rosy locus in Drosophila melanogaster. *Genetics.* 113:265-285.

Earnshaw, W. C., and N. Rothfield. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma (Berl.).* 91:313-321.

Earnshaw, W. C., N. Halligan, C. Cooke, and N. Rothfield. 1984. The kinetochore is part of the chromosome scaffold. *J. Cell Biol.* 98:352-357.

Earnshaw, W. C., B. Halligan, C. A. Cooke, M. M. S. Heck, and L. F. Liu. 1985. Topoisomerase II is a structural component of mitotic chromosome scaffolds. *J. Cell Biol.* 100:1706-1715.

Earnshaw, W. C., K. F. Sullivan, P. S. Machlin, C. A. Cooke, D. A. Kaiser, T. D. Pollard, N. F. Rothfield, and D. W. Cleveland. 1987. Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen. *J. Cell Biol.* 104:817-829.

Earnshaw, W. C., H. Ratrie, and G. Stetten. 1989. Visualization of centromere proteins CENP-B and CENP-C on a stable dicentric chromosome in cytological spreads. *Chromosoma (Berl.).* 98:1-12.

Earnshaw, W. C., R. L. Bernat, C. A. Cooke, and N. F. Rothfield. 1992. The role of the centromere/kinetochore in cell cycle control. *Cold Spring Harbor Symp. Quant. Biol.* In press.

Ellis, L., E. Clauser, D. O. Morgan, M. Edery, R. A. Roth, and W. J. Rutter. 1986. Replacement of insulin receptor tyrosine residues 1162 and 1163 compromises insulin-stimulated kinase activity and uptake of 2-deoxyglucose. *Cell.* 45:721-732.

Finch, J. T., and A. Klug. 1976. Solenoidal model for superstructure in chromatin. *Proc. Natl. Acad. Sci. USA.* 73:1897-1901.

Henikoff, S. 1981. Position-effect variegation and chromosome structure of a heat shock puff in Drosophila. *Chromosoma (Berl.).* 83:381-393.

Herskowitz, I. 1987. Functional inactivation of genes by dominant negative mutations. *Nature (Lond.).* 329:219-222.

Jabs, E. W., C. A. Goble, and G. R. Cutting. 1989. Macromolecular organization of human centromeric regions reveals high-frequency, polymorphic macro DNA repeats. *Proc. Natl. Acad. Sci. USA.* 86:202-206.

Jokelainen, P. T. 1967. The ultrastructure and spatial organization of the metaphase kinetochore in mitotic rat cells. *J. Ultrastruct. Res.* 19:19-44.

Joly, J. C., G. Flynn, and D. L. Purich. 1989. The microtubule-binding fragment of microtubule-associated protein-2: location of the protease-accessible site and identification of an assembly-promoting peptide. *J. Cell Biol.* 109:2289-2294.

Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith. 1984. A short amino acid sequence able to specify nuclear location. *Cell.* 39:499-509.

Kreis, T. E. 1986. Microinjected antibodies against the cytoplasmic domain of

vesicular stomatitis virus glycoprotein block its transport to the cell surface. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:931-941.

Laemmli, U. K., S. M. Cheng, K. W. Adolph, J. R. Paulson, J. A. Brown, and W. R. Baumbach. 1978. Metaphase chromosome structure: the role of nonhistone proteins. *Cold Spring Harbor Sym. Quant. Biol.* 423:351-360.

Luykx, P. 1965. The structure of the kinetochore in meiosis and mitosis in Urechis eggs. *Exp. Cell Res.* 39:643-657.

Machamer, C. E., and J. K. Rose. 1987. A specific transmembrane domain of a coronavirus E1 glycoprotein is required for its retention in the Golgi region. *J. Cell Biol.* 105:1205-1214.

Masumoto, H., H. Masukata, Y. Muro, N. Nozaki, and T. Okazaki. 1989. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J. Cell Biol.* 109:1963-1973.

Murre, C., P. S. McCaw, and D. Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell.* 56:777-783.

Musich, P. R., F. L. Brown, and J. J. Maio. 1982. Nucleosome phasing and micrococcal nuclease cleavage of African green monkey component a DNA. *Proc. Natl. Acad. Sci. USA.* 79:118-122.

Pfarr, C. M., M. Coue, P. M. Grissom, T. S. Hays, M. E. Porter, and J. R. McIntosh. 1990. Cytoplasmic dynein is localized to kinetochores during mitosis. *Nature (Lond.).* 345:263-265.

Pluta, A. F., C. A. Cooke, and W. C. Earnshaw. 1990. Structure of the human centromere at metaphase. *Trends Biochem. Sci.* 15:181-185.

Robbins, J., S. M. Dilworth, R. A. Laskey, and C. Dingwall. 1991. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell.* 64:615-623.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A laboratory manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY.

Simerly, C., R. Balczon, B. R. Brinkley, and G. Schatten. 1990. Microinjected kinetochore antibodies interfere with chromosome movement in meiotic and mitotic mouse embryos. *J. Cell Biol.* 111:1491-1504.

Steuer, E. R., L. Wordeman, T. A. Schroer, and M. P. Sheetz. 1990. Localization of cytoplasmic dynein to mitotic spindles and kinetochores. *Nature (Lond.).* 345:266-268.

Sullivan, K. F., and C. A. Glass. 1991. CENP-B is a highly conserved mammalian centromere protein with homology to the helix-loop-helix family of proteins. *Chromosoma (Berl.).* 100:360-370.

Waye, J. S., and H. Willard. 1986. Structure, organization, and sequence of alpha satellite DNA from human chromosome 17: evidence for evolution by unequal crossing-over and an ancestor pentamer repeat shared with the human X chromosome. *Mol. Cell. Biol.* 6:3156-3165.

Wevrick, R., and H. F. Willard. 1989. Long-range organization of tandem arrays of alpha satellite at the centromeres of human chromosomes: high frequency array length polymorphism and meiotic stability. *Proc. Natl. Acad. Sci. USA.* 86:9394-9398.

Widom, J., and A. Klug. 1985. Structure of the 300Å chromatin filament: x-ray diffraction from oriented samples. *Cell.* 43:207-213.

Willard, H. F., and J. S. Waye. 1987. Hierarchical order in chromosome-specific human alpha satellite DNA. *Trends Genet.* 3:192-198.

Wordeman, L., E. Steurer, M. Sheetz, and T. Mitchison. 1991. Chemical subdomains within the kinetochore domain of isolated CHO mitotic chromosomes. *J. Cell Biol.* 114:285-294.

Zhang, X.-Y., F. Fittler, and W. Horz. 1983. Eight different highly specific nucleosome phases on satellite DNA in the African green monkey. *Nucleic Acids Res.* 11:4275-4306.