

FZ

Biochemical Characterization of CBF3, an Essential DNA-binding Component of the Yeast Kinetochores

by

Adam Scott Grancell

B.S., Biological Sciences
Stanford University, 1991

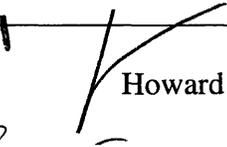
Submitted to the Department of Biology
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Biology

at the
Massachusetts Institute of Technology
June, 1998

© 1998 Massachusetts Institute of Technology
All rights reserved

Signature of Author:  _____
Department of Biology
May 29, 1998

Certified by:  _____
Peter K. Sorger
Howard and Linda Stern Assistant Professor of Biology
Thesis Supervisor

Accepted by:  _____
Frank Solomon
Professor of Biology
Chairman, Committee for Graduate Students

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY

AUG 04 1998 Science

LIBRARY

Biochemical Characterization of CBF3, an Essential DNA-binding Component of the Yeast Kinetochores

by

Adam Scott Grancell

Submitted to the Department of Biology on May 29, 1998, in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biology

Abstract

The kinetochore is a structure composed of proteins and centromeric DNA that attaches chromosomes to microtubules of the mitotic spindle. Correct assembly of kinetochores is essential for accurate chromosome segregation during mitosis and meiosis. The centromere of *Saccharomyces cerevisiae* contains 125 base pairs of DNA organized into three conserved regions, CDEI, CDEII, and CDEIII. CDEIII is bound by CBF3, a complex of four essential proteins, p110, p64, p58, and p23. Only p64 has a recognizable DNA-binding motif, a *GALA*-like zinc cluster domain. The oligomeric states, approximate shapes, and subunit-subunit interactions among CBF3 proteins have been characterized using gel filtration chromatography and velocity sedimentation. p110 is an elongated homodimer, p64 is a globular homodimer, p23 is monomeric, and p23 and p58 form a stable heterodimer in solution. p58 binds to each of the other CBF3 components individually. While there is no evidence that other CBF3 proteins can interact in the absence of p58, p58 mediates an interaction between p64 and p110. These results suggest that p58 plays a central role in the assembly of CBF3 onto centromeric DNA and allow us to build a scale model of the CBF3-DNA complex.

Thesis Supervisor: Peter K. Sorger

Title: Howard and Linda Stern Assistant Professor of Biology

Table of Contents

Abstract	2
Acknowledgements	4
Chapter 1	Introduction: Structural components of the yeast kinetochore	5
	<i>Figures</i>	24
Chapter 2	Physical characterization of individual CBF3 proteins and analysis of CBF3 subcomplex formation using hydrodynamic methods	32
	<i>Table and Figures</i>	53
Chapter 3	Final thoughts: A model for kinetochore assembly	75
	<i>Figure</i>	79
Appendix A	Physical domain mapping of p64 and p110 by partial proteolysis	81
	<i>Figures</i>	88
Appendix B	Chromosome movement: Kinetochores motor along (originally published in <i>Current Biology</i> 1998, 8: R382- R385)	94
Literature cited	99

Acknowledgements

When I began this journey almost seven years ago, I really believed that I would be doing it on my own. I was wrong. I would like to thank the many people who have supported me during my graduate work.

I owe a lot to Peter Sorger for taking me, as a fifth year graduate student, into his lab; for teaching me about yeast biology and protein biochemistry, about how to give a good presentation, and about carefully and convincingly writing up my work; and for being patient through the times when I was a pain in the ass.

I thank the members of my thesis committee, Steve Bell, Dean Dawson, Bob Sauer, Tony Sinskey, and Peter Sorger, for their valuable comments on my thesis and for their encouragement throughout the entire process of finishing up. I am especially grateful to Steve and Bob for conversation about my work over the course of my project.

I am grateful to Peter Sorger, Sheri Treadway, and Garth Wiens for critical reading of this thesis. I am particularly indebted to Peter, for reading multiple drafts very quickly. It would have been impossible for me to produce a readable thesis without their help.

I thank the entire Sorger lab for their questioning and commenting and teaching over the last two and a half years. In particular, I want to thank the CBF3 group, Chris Espelin, Ken Kaplan, and Iain Russell, for the reagents and methods and ideas that they shared. Ken has my eternal gratitude for setting up so many of the reagents I used every day and for occasionally making my bench look pretty clean (in comparison to his).

Finally, I would like to thank the following people for their support and encouragement during my time as a graduate student: my parents, grandparents, and siblings, especially my sister, Dannah; Liz Alcamo, Mark Borowsky, Matt and Monica Elrod-Erickson, Steve Escamilla, Carolyn Frenger, Jill Gillingham, Irwin Gross, Alan Grossman, Ken Kaplan, Mark Metzstein, Iain Russell, Bob Sauer, Frank Solomon, Peter Sorger, Gillian Stanfield, Sheri Treadway, and Anthony Van Niel. Their support gave me the strength and the confidence to accomplish what I have accomplished.

Chapter 1

Introduction: Structural components of the yeast kinetochore

During eukaryotic cell division, it is essential for correct chromosome segregation that each chromosome becomes correctly attached to the mitotic spindle. A correctly attached chromosome is one in which one sister chromatid is attached to microtubules emanating from one pole of the spindle, and the other chromatid is attached to microtubules from the opposite pole. The failure of chromosomes to attach properly to spindle microtubules can result in the generation of aneuploid daughter cells. While chromosome loss often leads to cell death, aneuploidy has also been associated with transformation and tumorigenesis (Cahill et al., 1998; Orr-Weaver and Weinberg, 1998).

The structure that mediates attachment of the chromosomes to the mitotic spindle is called the kinetochore, and it is present in one copy per sister chromatid in most organisms. In addition to its role in attaching chromosomes to microtubules, the kinetochore has been implicated in moving chromosomes along the walls of microtubules, mediating movement of chromosomes attached to the tips of dynamic microtubules, holding together sister chromatids in mitosis and homologous chromosomes during meiosis, and signaling to the cell cycle machinery that correct spindle attachment has occurred. *In vivo* assays for kinetochore function measure only chromosome loss, which could be a consequence of the failure of any of these kinetochore activities or of a defect in DNA synthesis.

The kinetochore is composed of proteins and centromeric DNA (CEN DNA). In humans and fruit flies, the size of the centromere is measured in megabases and hundreds of kilobases, respectively, and is composed primarily of repetitive DNA sequences (reviewed in Murphy and Karpen, 1998; Wiens and Sorger, 1998). Electron microscopy reveals the mammalian kinetochore to be a distinct three-layered “trilaminar” structure associated with 10-30 microtubules. Proteins that localize to this trilaminar structure include a histone H3 homolog (CENP-A), an alpha-satellite DNA repetitive element-binding protein (CENP-B), and a microtubule motor (CENP-E) (Cooke et al., 1990; Warburton et al., 1997; Yao et al., 1997). The analysis of these proteins has been limited by the difficulty of experimental manipulation in animal cells.

In contrast to higher eukaryotes, the yeast *Saccharomyces cerevisiae* provides a simple system for analysis of kinetochore structure and function. The yeast centromere contains only 125 base pairs of conserved DNA sequences, and many proteins that bind to the centromeric sequences have been identified and analyzed both biochemically and genetically. In contrast to the mammalian kinetochore, the yeast kinetochore does not form a distinct structure in the electron microscope. And, whereas the mammalian kinetochore contacts multiple microtubules, each yeast kinetochore associates with only one microtubule (Peterson and Ris, 1976; Winey et al., 1995). Nevertheless, the functions of kinetochores in yeast and higher eukaryotes appear to overlap, and some of the proteins suspected to be involved in kinetochore function in the different organisms share significant sequence homologies. Thus, *S. cerevisiae*, which is highly amenable to both biochemical and genetic analysis, provides an excellent model system for understanding the structure and function of the kinetochore.

The purpose of this chapter is to describe what is currently known about the structural components of the yeast kinetochore. While much recent work has investigated the molecular basis of sister chromatid cohesion (reviewed in Allshire, 1997) and mitotic checkpoint control (for reviews, see Nicklas, 1997; Straight, 1997), these topics will not be addressed here. And, although reference will be made to non-yeast systems, discussion will be confined to the biology of *S. cerevisiae* unless otherwise noted.

Assays for kinetochore function

The kinetochore is required for correct chromosome segregation, so defects in kinetochore function *in vivo* are manifested by increased rates of chromosome loss. Mutations in kinetochore structural proteins and regulatory factors, in mitotic spindle components, and in CEN DNA all increase chromosome loss. During mitosis, wild type yeast lose any one of the sixteen native chromosomes approximately once per 10^5 divisions (Hartwell et al., 1982). Chromosome loss is assayed by detecting the loss of a dominant marker present on a “test chromosome”. Various investigators have used different types of test chromosomes, including circular plasmids (minichromosomes), linear chromosome fragments, native yeast chromosomes, and yeast artificial chromosomes (YACs).

In the context of these assays, “chromosome loss” is a phenotypic, not a mechanistic, term. That is, the loss of a dominant marker can occur by many different mechanisms, including failure to replicate, nondisjunction, premature disjunction, failure of kinetochore-microtubule attachment, and actual loss of DNA. The “sectoring assay” can distinguish among some of these abnormalities by differentiating between cells

containing different numbers of test chromosomes (Hieter et al., 1985). In this assay, a genomic *ade2* mutation turns colonies red due to a defect in the adenine biosynthesis pathway. Test chromosomes contain a single copy of *SUP11*, a tRNA_{ochre} gene that suppresses the *ade2* mutation and turns diploid yeast colonies pink or white, depending on whether the cells contain one or two copies of the suppressor, respectively. Knowing the number of chromosomes in the daughter cells can help distinguish between mechanisms of missegregation. For example, a 1:0 division, in which one daughter receives one copy of a chromosome and the other receives no copies, is much more likely to be caused by actual chromosome loss or failure to replicate than by nondisjunction.

Centromeric DNA

Isolation of yeast centromeres

The first centromeric DNA isolated from *S. cerevisiae* was the centromere of chromosome III (*CEN3*) (Clarke and Carbon, 1980a). *CEN3* was identified based on its ability to convey stable mitotic segregation to a circular plasmid containing an autonomous replication sequence (ARS). Normally, ARS plasmids, which are present in 50-100 copies per cell, are lost rapidly when cells are grown in conditions that do not select for an auxotrophic marker present on the plasmid (Clarke and Carbon, 1980b). This loss occurs because ARS plasmids do not engage the mitotic spindle and as a result segregate disproportionately to the mother cell during cell division (Murray and Szostak, 1983). A DNA fragment containing the *CDC10* gene and adjacent sequences was found to increase transmission frequency of an ARS plasmid greater than 90-fold after ten generations of non-selective growth (Clarke and Carbon, 1980a; Clarke and Carbon, 1980b). This plasmid was estimated to be present in approximately one copy per cell,

and it segregated correctly during meiosis. *CDC10* was known to be closely linked to the centromere of chromosome III by high resolution linkage analysis (Mortimer and Schild, 1980), and subcloning of the *CDC10*-containing DNA fragment identified a region of DNA that, when present in *cis*, conveyed mitotic and meiotic stability to an ARS plasmid. Removal of this DNA sequence from the endogenous chromosome III caused the chromosome to be lost rapidly during cell division, and the cloned region could be substituted for a genomic centromere without affecting chromosome segregation (Clarke and Carbon, 1983). These last experiments demonstrated that the cloned fragment contained sequences necessary and sufficient for centromere function and therefore contained an actual centromere.

The cloning of additional yeast centromeres led to the recognition of conserved sequence regions within CEN DNA. Several centromeres were isolated by cloning yeast biosynthetic genes closely linked to centromeres, retrieving adjacent DNA sequences by overlap hybridization (“chromosome walking”), and then identifying the CEN DNA by its ability to stabilize the transmission of an ARS plasmid (Fitzgerald-Hayes et al., 1982; Panzeri and Philippsen, 1982; Stinchcomb et al., 1982; Mann and Davis, 1986). Other centromeres were identified in genetic screens looking for DNA fragments that either stabilized an ARS plasmid (Hsiao and Carbon, 1981) or reduced ARS plasmid copy number (Hieter et al., 1985) and by the sequencing of the entire yeast genome.

Nucleotide sequence analysis of the centromeres showed that all functional centromeres contained an approximately 125 base pair region of DNA composed of three adjacent conserved DNA elements (CDEs), CDEI, CDEII, and CDEIII (Fitzgerald-Hayes et al., 1982; Hieter et al., 1985) (Figure 1.1). By convention, the CDEI proximal side of the

centromere will be referred to as the left and the CDEIII proximal side will be referred to as the right.

Structure-function analysis of conserved centromere sequences

The three regions of the centromere, when placed in the context of a larger piece of DNA, are necessary and sufficient for complete mitotic centromere function (Cottarel et al., 1989). As shown in Figure 1.2A, CDEI is an eight base pair partial palindrome consisting of the sequence RTCACRTG, where R is a purine. Deletion of CDEI increases the rate of chromosome loss during mitosis 10 to 30-fold (Niedenthal et al., 1991). The six nucleotide palindrome on the right side of CDEI is most important for function, since mutations in these nucleotides can have the same effect on chromosome loss as a complete CDEI deletion (Niedenthal et al., 1991). The high conservation of CDEI sequences suggests that CDEI interacts with a site-specific DNA-binding protein.

The precise sequence of CDEII varies significantly among the sixteen yeast centromeres, as shown in Figure 1.3. CDEIIs range in size from 78 to 87 nucleotides and are distinguished by a high content (greater than 85%) of A+T residues. The arrangement of nucleotides is not entirely random, consisting primarily of tracks of identical residues up to eight bases long. Removing larger and larger amounts of CDEII sequence progressively increases chromosome loss rates (Cumberledge and Carbon, 1987; Gaudet and Fitzgerald-Hayes, 1987), and completely deleting CDEII destroys the centromere (Gaudet and Fitzgerald-Hayes, 1987). Centromere function is also sensitive to increasing the length of CDEII: doubling the length of CDEII by adding a second CDEII sequence in tandem increases chromosome loss 100-fold (Gaudet and Fitzgerald-Hayes, 1987). The importance of the A+T content and the precise arrangement of the

sequence is uncertain. Substituting half of CDEII with DNA that is not A+T-rich or with DNA that is A+T-rich but not organized into tracks reduces centromere function 40- or 4-fold, respectively (Cumberledge and Carbon, 1987). However, replacing the entire CDEII region with non-CDEII A+T-rich sequence destroys the centromere (Heus et al., 1994). The lack of exact conservation and relative insensitivity to some large sequence changes suggest that if CDEII is bound by sequence-specific, as opposed to structure-specific, DNA binding proteins, the recognition sites for these proteins are fairly general.

Unlike CDEI and CDEII, CDEIII contains individual nucleotides that are essential for centromere function. CDEIII, shown in Figure 1.2A, is roughly an imperfect palindrome centered around a cytosine residue. The effect of altering conserved nucleotides varies non-randomly throughout CDEIII and is summarized in Figure 1.2B. Mutations in the central CCG are the most severe, followed by mutations in bases adjacent to CCG, and then by mutations in the sequences near the edges of CDEIII (Hegemann et al., 1988; Jehn et al., 1991). Although CDEIII is somewhat palindromic, mutations on different sides of the palindrome have quantitatively different effects. Within the central TTCCGAA sequence, mutations on the right side are more severe than mutations on the left. Mutations on the left edge of CDEIII are more severe than those on the right edge. This asymmetry is important for CDEIII function, since centromeres in which CDEIII is inverted relative to CDEI and CDEII are inactive (Murphy et al., 1991). The high conservation of CDEIII suggests that it binds site-specific DNA binding proteins, and the asymmetry demonstrates that the proteins have a defined orientation, suggesting that CDEIII-bound proteins participate in specific interactions with other kinetochore components.

The integrity of the centromere is also required for proper chromosome segregation during meiosis. Alterations in both CDEI and CDEII cause defects in meiosis that are more severe than the mitotic defects associated with these mutations (Sears et al., 1995). These changes primarily increase premature separation of sister chromatids during meiosis I and nondisjunction during meiosis II, but they have less of an effect on meiosis I nondisjunction. Deletions of portions of CDEII have the most severe effects on meiotic chromosome segregation, including a seventeen-fold increase in meiosis II nondisjunction. The relationship between the meiotic and mitotic defects is unclear, although the observation that centromere mutants have relatively minor increases in meiosis I nondisjunction is consistent with there being fundamental differences between the kinetochores found in homologous chromosomes and in sister chromatids.

Kinetochores proteins

The chromatin structure of the centromere is distinct from that of the surrounding DNA. In DNA isolated from yeast cells, the ~125 base pair centromere as well as sequences to its left and right are protected from nuclease digestion, suggesting that a 150-250 base pair region is bound up in a large nucleoprotein complex (Bloom and Carbon, 1982; Funk et al., 1989). It has been postulated that this 150-250 base pair region and the proteins bound to it were the functional equivalent of the mammalian trilaminar kinetochore observed in the electron microscope (Bloom and Carbon, 1982). Alterations in the nuclease resistance of this sequence in the presence of centromere mutations supported this hypothesis (Saunders et al., 1988), and suggested that chromosome missegregation associated with these mutations was due to the failure of kinetochore proteins to assemble properly at the centromere. Subsequently, Cbf1p and

the protein complex CBF3, which bind in a sequence-specific manner to CDEI and CDEIII, respectively, were isolated, and genetic analysis showed that these proteins were indeed required for kinetochore function. Other potential kinetochore proteins have been identified either biochemically or genetically, but none of them have been shown to bind to CDEII. In this section, the isolation and initial characterization of proteins demonstrated and suspected to function at the kinetochore will be described.

Cbf1p

Cbf1p is a helix-loop-helix protein that binds to CDEI DNA. Alternatively called CBP-1 (Cai and Davis, 1989), CP1 (Bram and Kornberg, 1987; Baker et al., 1989), and CPF1 (Jiang and Philippsen, 1989), Cbf1p was identified based on its ability to bind CDEI sequences at the centromere. Mitotic chromosome loss rates are increased up to 10-fold by deleting or mutating Cbf1p (Mellor et al., 1991; Foreman and Davis, 1993), demonstrating that Cbf1p function is important for centromere function. Several lines of evidence suggest that binding of Cbf1p to CDEI is important for the *in vivo* function of Cbf1p. First, a CEN plasmid containing a CDEI deletion is lost at the same rate in wild type and *CBF1* deletion backgrounds, suggesting that Cbf1p mediates the function of CDEI (Baker and Masison, 1990). Second, mutations in CDEI that decrease its ability to bind Cbf1p *in vitro* increase the loss rates of plasmids or linear chromosome fragments bearing those mutations (Baker et al., 1989; Cai and Davis, 1989; Wilmen et al., 1994). Conversely, point mutations in Cbf1p that decrease its ability to bind to CDEI increase chromosome loss *in vivo* (Foreman and Davis, 1993). Third, *in vivo* footprinting experiments demonstrate that CDEI is protected from dimethyl sulfate methylation in a

wild type strain but not in a strain deleted for *CBF1* (Densmore et al., 1991), suggesting that Cbf1p binds to CDEI in the cell.

In addition to its role in chromosome segregation, Cbf1p functions as a transcriptional activator for a variety of genes involved in metabolism. Cells with *CBF1* deletions display methionine auxotrophy and slow growth which cannot be accounted for by increased chromosome loss alone (Baker and Masison, 1990). The CDEI sequence is found in many upstream activation sequences (UASs), including *MET16*, *MET25*, and *GAL2* (Bram and Kornberg, 1987; Kent et al., 1994). Cbf1p assembles with the products of the *MET28* and *MET4* genes into a complex that binds CDEI in the UAS of *MET16* (Kuras et al., 1997). It is believed that the Cbf1p/Met28p/Met4p complex clears nucleosomes from the DNA surrounding its binding site, making these regions more accessible to general transcription factors (Kent et al., 1994; O'connell et al., 1995). The relationship of this chromatin remodeling function at transcription sites to the role of Cbf1p at kinetochores is unclear, especially because point mutations in Cbf1p can separate methionine auxotrophy from chromosome loss phenotypes (Foreman and Davis, 1993).

CBF3

CBF3 was initially identified as a set of proteins in yeast nuclear extract that bound specifically to CDEIII DNA *in vitro* (Ng and Carbon, 1987; Lechner and Carbon, 1991). This complex includes four proteins, p110, p64, p58, and p23, encoded by the genes *NDC10/CTF14/CBF2* (Doheny et al., 1993; Goh and Kilmartin, 1993; Jiang et al., 1993), *CEP3/CBF3b* (Lechner, 1994; Strunnikov et al., 1995), *CTF13* (Doheny et al., 1993), and *SKP1/CBF3d* (Connelly and Hieter, 1996; Stemmann and Lechner, 1996),

respectively. These genes are all essential for growth, and mutations in them are associated with increased chromosome loss (Doheny et al., 1993; Goh and Kilmartin, 1993; Jiang et al., 1993; Lechner, 1994; Strunnikov et al., 1995; Connelly and Hieter, 1996; Stemmann and Lechner, 1996). None of the CBF3 proteins binds to CDEIII on its own (Stemmann and Lechner, 1996; Espelin et al., 1997), but the combination of all four is necessary and sufficient for binding to CDEIII *in vitro* (Espelin et al., 1997; Kaplan et al., 1997). Extracts prepared from cells containing CBF3 protein mutants fail to support *in vitro* assembly of CBF3 onto DNA (Doheny et al., 1993; Sorger et al., 1995; Kaplan et al., 1997), and mutated CDEIII sequences that compromise centromere function *in vivo* also decrease or inhibit binding of CBF3 to DNA *in vitro* and *in vivo* (Ng and Carbon, 1987; Lechner and Carbon, 1991; Sorger et al., 1995; Espelin et al., 1997; Meluh and Koshland, 1997). These results strongly suggest that the binding of CBF3 to CDEIII is required for kinetochore function.

A large collection of evidence supports the notion that CBF3 initiates assembly of the kinetochore on centromeric DNA. First, beads coated with CDEIII DNA can bind to microtubules *in vitro* when incubated with yeast nuclear extract but not when incubated with purified CBF3 alone (Sorger et al., 1994). This result suggests that CBF3, in addition to binding CDEIII, interacts with unidentified microtubule binding proteins. Second, CDEIII is necessary to recruit both Cbf1p and another kinetochore component, Mif2p, to the centromeres *in vivo*, although neither protein binds directly to CDEIII (Meluh and Koshland, 1997). Third, mutations in CDEIII that disrupt CBF3 binding *in vitro* completely eliminate the nuclease protection of centromeric DNA isolated from cells (Saunders et al., 1988). In contrast, mutations in CDEI or CDEII only alter the size

of the protected region. This result suggests that binding of CBF3 to CDEIII is necessary for assembly of the remaining kinetochore components at the centromere. Based on these findings, a reasonable working model is that CBF3 initiates kinetochore formation. However, confirmation of this model will require identifying other kinetochore components and establishing that CBF3 must bind to CDEIII before the components can assemble.

Only one of the CBF3 proteins, p64, has a recognizable DNA-binding motif. p64 is a member of a family of proteins characterized by an approximately 30 amino acid $Zn(II)_2Cys_6$ zinc cluster DNA binding domain. Most of the zinc cluster proteins, including Gal4p, Ppr1p, and Hap1p, are fungal transcription factors that bind alone as homodimers to either direct or inverted repeats of the sequence CCG (Marmorstein et al., 1992; Marmorstein and Harrison, 1994; Zhang and Guarente, 1994). In contrast, p64 binds to DNA only in the presence of the remaining CBF3 proteins, and CDEIII contains only a single CCG. Protein-DNA crosslinking studies show that p64 contacts two regions of CDEIII DNA, including the essential CCG sequence (Espelin et al., 1997). Mutating residues of the p64 zinc cluster that correspond to amino acids important for zinc cluster DNA binding in Gal4p and Ppr1p increases chromosome loss in yeast (Lechner, 1994). These results support the notions that the zinc cluster of p64 contacts the CCG of CDEIII and that this binding is critical for kinetochore function. It seems reasonable that, because there is only a single CCG in CDEIII, p64 requires interaction with other proteins to enable zinc cluster DNA binding. However, the precise protein-protein interactions that allow p64 to bind DNA are not yet known.

Organization of CBF3 on DNA

While the activities of all four CBF3 proteins are required for binding of CBF3 to CDEIII, only three proteins have actually been shown to contact DNA (Espelin et al., 1997). The results of protein-DNA crosslinking studies are summarized in Figure 1.4. As discussed above, p64 contacts two sites in CDEIII, the CCG and 3-4 bases approximately one helical turn to the left of CCG. p58 contacts a trinucleotide region that falls in between the two p64-interacting sites, suggesting that p58 and p64 sit on opposite sides of the DNA helix directly across from one another. Unlike p58 and p64, p110 does not interact with discrete regions, contacting nucleotides throughout CDEIII and 25 base pairs to the right. The presence of additional nonspecific DNA to the right enables formation of a lower-mobility, “extended” CBF3 complex in which additional p110 protein interacts with the new DNA sequences. These results demonstrate that the binding of CBF3 to CDEIII is asymmetric, which may explain why inverting CDEIII relative to CDEI and CDEII inactivates the centromere.

Regulation of CBF3 assembly

Calculations based on the yield of CBF3 purified from yeast show that each cell contains only as many CBF3 complexes as centromeres, suggesting that the cell regulates CBF3 levels (Lechner and Carbon, 1991). A possible mechanism for this regulation is demonstrated by the finding that functional p58 levels may be regulated by two opposed mechanisms (Kaplan et al., 1997). The first mechanism is post-translational activation of p58. p58 expressed on its own in a recombinant system is incapable of participating in CBF3-DNA complex formation. However, when coexpressed with p23, p58 is functional in CBF3 DNA binding. This “activation” event probably involves the phosphorylation of p58 by an unidentified kinase. That similar activation occurs in yeast was demonstrated

by the finding that activated p58 that has been coexpressed with p23 and then purified away from the p23 can complement defects in the *in vitro* DNA binding of extracts from either p58 or p23 mutant strains. The second mechanism for controlling p58 is regulated degradation. p58 overexpressed in yeast has a half life of approximately 15 minutes. In contrast, p64 and p110 have half lives greater than five hours. Mutations in components of the 20S proteasome, the proteolytic machinery required for degrading ubiquitin-tagged proteins, increases the half life of p58. In addition, p58 may be covalently modified before its degradation. These results suggest that p58 is specifically tagged by ubiquitin or a ubiquitin homolog and targeted for degradation.

Why would it be important to control the levels of active CBF3 proteins? One possibility is that low levels of CBF3 enhance the specificity of CBF3 DNA binding. Specificity of kinetochore assembly is important, because extra kinetochores can lead to attachment of a chromatid to both spindle poles and subsequent fragmentation during anaphase (Mann and Davis, 1983). Alternatively, because there are so few centromeres in the cell, it may be necessary to limit free CBF3 protein levels to prevent titrating out other kinetochore components. In either scenario, chromosome loss would be increased by increasing CBF3 levels.

The function of p23

In addition to its role in activating p58 for kinetochore function, p23 is also required for the degradation of cell cycle regulating proteins. p23 is a component of SCF, a complex that ligates ubiquitin moieties onto proteins to target them for degradation (Feldman et al., 1997; Skowyra et al., 1997). There are multiple forms of SCF, and they all include p23 and the protein Cdc53p (Skowyra et al., 1997). In

addition, the different SCF complexes contain one of many proteins that are responsible for specifically recognizing the target for ubiquitination. These targeting proteins each contain an F-box, a region required for association with p23 (reviewed in Krek, 1998). The relationship between the functions of p23 in kinetochore assembly and in ubiquitin-mediated degradation are unclear. p58 contains an F-box sequence which mediates its interaction with p23 (I. Russell, unpublished observation). However, mutations in p23 can separate its different functions (Connelly and Hieter, 1996). One intriguing possibility is that the p23-p58 interaction concomitant with p23-dependent activation of p58 recruits p23 to the kinetochore to serve a function more closely related to its role in degradation. Identification of other F-box proteins and determining whether any of them play roles at the kinetochore may be useful in exploring the functions of p23.

Other kinetochore proteins

Mif2p

MIF2 was initially identified as a gene fragment that caused aberrant chromosome segregation when expressed at high levels (Brown et al., 1993). Mutations in *MIF2* have a synergistic effect on chromosome loss when present in combination with *CDEI* mutations, and are synthetically lethal in combination with *CBF1*, *NDC10* and *CEP3* mutations (Meluh and Koshland, 1995). *In vivo*, Mif2p associates with centromeric DNA in a manner dependent on functional *CDEIII* (Meluh and Koshland, 1997), demonstrating that Mif2p is an actual component of the kinetochore. While this result suggests that Mif2p interacts with CBF3, the observation that overexpressing Cbf1p but not p110 or p58 can partially rescue *mif2* viability defects (Meluh and Koshland, 1995) supports the notion that Cbf1p and Mif2p contribute to the same function at the kinetochore.

Sequence analysis of *MIF2* reveals two homologies that may be relevant to kinetochore function. First, *MIF2* shares homology with regions important for the function of the protein CENP-C, a component of the inner plate of the mammalian kinetochore (Saitoh et al., 1992; Meluh and Koshland, 1995). Second, Mif2p contains an eight amino acid “A-T hook” motif that allows HMG-I(Y) proteins to bind in the minor groove of A+T-rich DNA (Reeves and Nissen, 1990). Although this homology suggests that Mif2p might bind to the A+T-rich CDEII sequence, deletion of the A-T hook region of Mif2p results in a protein that is still functional *in vivo* (Jaffe, 1996), demonstrating that this region is not required for Mif2p function.

Cse4p

CSE4 was identified in a screen for mutations that increase the loss rate of a chromosome bearing a centromere with a partially deleted CDEII (Stoler et al., 1995). Sequence analysis demonstrated that Cse4p shares a region of significant homology with histone H3 and with CENP-A, a histone H3 homolog found specifically in the inner plate region of active mammalian centromeres (Warburton et al., 1997). This homology is in the histone fold, a domain that participates in both protein-protein and protein-DNA interactions in the nucleosome (Luger et al., 1997). In addition to histone sequence homology, Cse4p coelutes with histones from DNA (Stoler et al., 1995). These results suggest that Cse4p is a kinetochore-specific histone H3 variant that assembles into nucleosomes either on or surrounding centromeric DNA. Additional support for this notion comes from the finding that high copy *CSE4* suppresses growth defects of a histone H4 mutant that is defective in chromosome segregation (Smith et al., 1996).

Chromatin

In addition to the possibility that Cse4p is a component of kinetochore-specific chromatin, several other results suggest that the chromatin structure in or around the centromere is important for kinetochore function. Mutations in Spt4p, a protein believed to be involved in chromatin remodeling during transcription, and in the histone H4 can increase chromosome loss rates (Basrai et al., 1996; Smith et al., 1996). Depletion of histones from yeast alters the regions of centromeric DNA resistant to nuclease digestions, specifically increasing the sensitivity of CDEII (Saunders et al., 1990). While these results do not distinguish between direct and indirect effects of chromatin structure on the kinetochore, they raise the possibility that DNA conformation or interactions with nucleosomes influences kinetochore assembly or stability.

Microtubule motors

It is expected that the yeast kinetochore will contain at least one motor protein, both because of the microtubule-based functions of the kinetochore and because motor proteins localize to the kinetochores of higher cells (Pfarr et al., 1990; Steuer et al., 1990; Yen et al., 1992; Wordeman and Mitchison, 1995; Walczak et al., 1996). At least one mammalian motor protein, the kinesin-like motor CENP-E, has been shown to be required for microtubule-mediated chromosome movement *in vitro* and *in vivo* (Schaar et al., 1997; Wood et al., 1997). With the complete sequencing of the *S. cerevisiae* genome, it is known that yeast contain only six kinesin-related proteins, and some of them have been shown to be required for events during mitosis (Hoyt et al., 1992). One yeast motor protein, Kar3p, copurifies with CBF3 on CDEIII DNA (Middleton and Carbon, 1994). Kar3p has *in vitro* microtubule motor activity and can mediate the binding of latex beads coated with CEN DNA to microtubules in the presence of CBF3 (Middleton and Carbon,

1994). However, the non-specific DNA binding activity of Kar3p may account for some of this bead binding activity. Deletions of *KAR3* lead to defects in nuclear fusion during mating (Meluh and Rose, 1990) and in synaptonemal complex formation during meiosis (Bascom-Slack and Dawson, 1997), but there is no genetic evidence that it functions at the kinetochore.

Cbf5p

Cbf5p was identified as a protein from yeast extracts that bound with low affinity to yeast centromeric DNA (Jiang et al., 1993). Cbf5p binds to microtubules *in vitro* and contains homology to known microtubule associated proteins (Jiang et al., 1993). *CBF5* is an essential gene that interacts genetically with both *NDC10* and *MCK1*, a gene encoding a serine/threonine kinase believed to play a role in chromosome segregation (Jiang et al., 1995). Recent experiments have shown that Cbf5p is required for the biosynthesis and covalent modification of ribosomal RNA (Lafontaine et al., 1998). Although the role of Cbf5p in rRNA metabolism is not inconsistent with the protein also functioning in chromosome segregation, the evidence that Cbf5p acts at the kinetochore is inconclusive.

Conclusions

While the past twenty years have seen huge advances in our understanding of centromeric DNA and DNA-binding components of the kinetochore, relatively little is known about the proteins that bind to microtubules or about the kinetochore assembly pathway. Identification of microtubule binding proteins will require further genetic analysis, including development of screens specific for isolating proteins functioning at the interface between microtubules and CEN-binding proteins, and the biochemical

isolation of proteins that interact physically with known kinetochore components. At this point, assembly questions revolve around CBF3, including how the various CBF3 proteins and DNA sequences contribute to CBF3 complex formation and how this assembly is regulated. Understanding the CBF3-DNA structure, which appears to play a central role in kinetochore function, is also expected to give insight into how other large nucleoprotein complexes assemble at specific sites in the genome.

Figure 1.1. Organization of the yeast centromere. A 125 base pair region of the centromere is necessary and sufficient for full function during mitosis and meiosis. The centromere is composed of three adjacent conserved DNA elements, CDEI, CDEII, and CDEIII. Arrows indicate palindromic sequences.

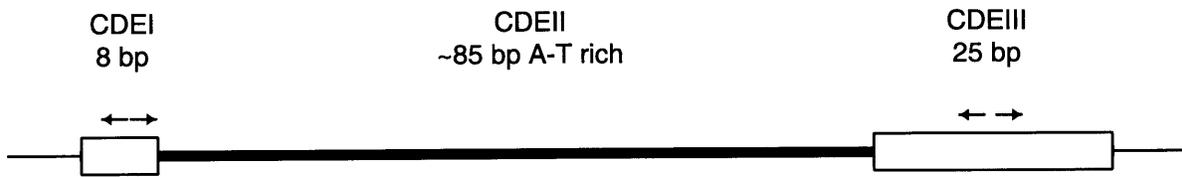


Figure 1.2. Conserved centromere sequences CDEI and CDEIII. (A) The sequences of CDEI and CDEIII regions from all sixteen yeast centromeres are shown. Consensus residues and the number of times out of sixteen each residue occurs are indicated below the sequences. In the CDEI consensus, R = purine. Sequences were retrieved from the Saccharomyces Genome Database (Cherry et al.). (B) A summary of the effects of CDEIII point mutations on centromere function *in vivo*. Chromosome fragments bearing point mutations in the CDEIII region of *CEN6* were assayed for loss. For each position, the effect of the most severe point mutation is shown. No data has been collected for bases marked with asterisks. Results are summarized from data presented in Jehn et al. (1991).

A

	CDEI	CDEIII
<i>CEN1</i>	G T C A C A T G	T G T T T T T G T T T T C C G A A G C A G T C A A
<i>CEN2</i>	A T C A T G T G	T G T T T T T G T T T T C C G A A A A A G A A A A
<i>CEN3</i>	G T C A C A T G	T G T A T T T G A T T T C C G A A A G T T A A A A
<i>CEN4</i>	G T C A C A T G	T G T T T A T G A T T A C C G A A A C A T A A A A
<i>CEN5</i>	A T C A C G T G	A G T A T T A G A T T T C C G A A A A G A A A A A
<i>CEN6</i>	A T C A C G T G	A G T T T T T G T T T T C C G A A G A T G T A A A
<i>CEN7</i>	A T C A C G T G	T G T T T T T G C C T T C C G A A A A G A A A A T
<i>CEN8</i>	A T C A C A T G	G G G T T T T G T G T T C C G A A C T T A G A A A
<i>CEN9</i>	T T C A C G T G	T G G T T T T G T T T T C C G A A A T G T T T T T
<i>CEN10</i>	A T C A C G T G	T G T T T A T G A T T T C C G A A C C T A A A T A
<i>CEN11</i>	G T C A C A T G	T G T T C A T G A T T T C C G A A C G T A T A A A
<i>CEN12</i>	A T C A C G T G	T G T A T T T G T T A T C C G A A C A A T A A A A
<i>CEN13</i>	A T C A C A T G	T G T G T A T G C G T T C C G A A C T T T A A A T
<i>CEN14</i>	G T C A C G T G	T G T A T T T G T C T T C C G A A A A G T A A A A
<i>CEN15</i>	A T C A C G T G	T G T A T A T G A C T T C C G A A A A A T A T A T
<i>CEN16</i>	A T C A C A T G	T G G T T A A G A T T T C C G A A A A T A G A A A
	16 16 16 16 16 16 16 16	13 16 13 - 15 - 14 16 - - 15 15 16 16 16 16 16 - - - - - 13 14 12
Consensus	R T C A C R T G	T G T - T ^A / _T T G - - T T C C G A A - - - - - A A A

B

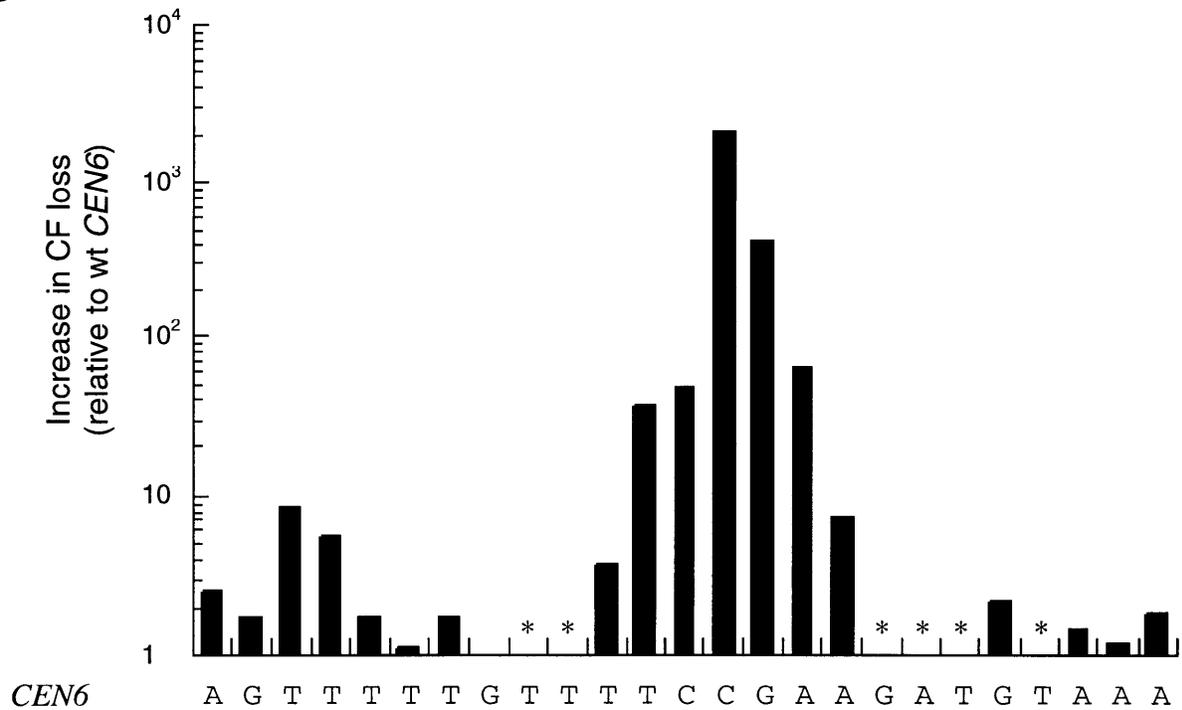


Figure 1.3. CDEII nucleotide sequences. The DNA sequences of all sixteen CDEII regions are presented, and the length and A+T content of each are indicated on the right. Although no precise sequences are conserved, tracks of identical bases are spread throughout each CDEII.

CDEII

		length	%A+T
<i>CEN1</i>	ACATAATAATAAATAATTTTAAAAATATAAAATATTTTAAATAGTTTTTAAATATTTTACAGTTTATTTTTTAAAATTTATTTATA	85	95
<i>CEN2</i>	ACTTATTTATTTAATTTATTTAAGTAAAAAAGATTTTCTATTTAAATTTATTAATTAATTTTTTTTTCTTAAATAATTTATTTTA	84	94
<i>CEN3</i>	ATGATATTTGATTTTATTATATTTTTTAAAAAAGTAAAAAATAAAAAGTAGTTTATTTTTTAAAAAATAAAATTTAAAAATATTAG	84	93
<i>CEN4</i>	CTTATAATCAACTTTTTTAAAAATTTAAAATACTTTTTTATTTTTTATTTTTTAAACATAAATGAAATAATTTATTTAT	78	92
<i>CEN5</i>	CTTTTTAAAAAATATAAATTTAATTTCAATTTCTATTTCAATATTTATTAAATAAAAAATTTGAAAAATATATAAAAAATTGTAGC	85	92
<i>CEN6</i>	CTATAAAAAATAATTATAAATTTAAATTTTTTAAATATAAATATATAAATTTAAAAATAGAAAGTAAAAAAGAAATTAAGAAAAAAT	85	94
<i>CEN7</i>	TTATATTTACTATATAAAAAATTCATAAATAAAAAGTTAGAAGATAAAAAATTATATTATACATATTTTTATTTTTATTATAATTTT	86	93
<i>CEN8</i>	ACTAATAATTCTTTTAATTTTAAATTTAATTTAATAAAAATTAATAAATATATATACTAAATTGTTTATTAAAAATGATTAAACATT	85	93
<i>CEN9</i>	AAAATTTTTTATATTTTTTAAATTTAAATTTTTTATAAATATTATAAATTTATTAATATTGATATTTAAAATTTAAAAACAATTTATTAA	84	98
<i>CEN10</i>	TTAAATAATTAATTTACTTTTAAAATTTATTTTTTAAATATAAAATATTTATTCTTTTTATTTTAAAAATAAAAAACACAAAAAACA	86	94
<i>CEN11</i>	ATAAAAACATATTTAAAATTTTAAAAAATTAATTTTCAAAATAAATTTATTATATTTTTTTAATTACATAATCATAAAAATAAA	85	95
<i>CEN12</i>	TAATAAATATTATTAAAAAGTTTATTAAAAATAAATAAATTTAAATTTACTATTTTTTAAATAAGTTTATTTTTTTAATAACACTAT	87	94
<i>CEN13</i>	ACTACCTAACAAAATATTTATTTTTCTTTTTTAAATTTTGAAAATACTAAAATATTTTTTGTTGTTTTTTGAAAAAAGGATTTTTAA	86	86
<i>CEN14</i>	CAGCTTTTTTAAAAATATTTTAAAACATTTTAAAAAATATACATTTTTTTTTATTATTTTTTTTATATATTAATGTTAAAAATTTATTTA	85	93
<i>CEN15</i>	AAC TTATTTTGCA TTTAAAAAAGTAAAAACTATTTGCTAAAATATATTTTTTTAAATTTTTAAAAATAATGTTTTAATTTATTTAA	86	91
<i>CEN16</i>	ATATATTTTTTATTTTTTAAATTTTTTTTAAATTTATAAAAAATAATTTTTTTCTTTAAATTTAAACAAAAATAAAAAATGTTTTTTGT	84	95

Figure 1.4. Regions of *CEN3* contacted by CBF3 proteins. The 89 base pair region contacted by CBF3 proteins *in vitro* is presented. p58 and p64 contact discrete sequences (shaded regions) within CDEIII (dashed box), and p110 makes contacts along an 89 base pair stretch of DNA. The 56 base pair "core" is the minimal sequence required for CBF3-DNA complex formation and corresponds to the region protected from nuclease digestion by CBF3 *in vitro*. Results are summarized from Espelin et al. (1997).



Chapter 2

Physical characterization of individual CBF3 proteins and analysis of CBF3 subcomplex formation using hydrodynamic methods

Introduction

In Chapter 1, I argued that CBF3 is an essential DNA binding component of the yeast kinetochore and may function to initiate kinetochore assembly. The four protein components of CBF3, p110, p64, p58, and p23, are necessary and sufficient for the binding of CBF3 to DNA *in vitro* (Kaplan et al., 1997). This finding allowed Espelin, et al. (1997) to determine, using recombinant CBF3 proteins in protein-DNA crosslinking experiments, that p58, p64 and p110 directly contact DNA. In this chapter, I describe experiments that explore interactions among the protein components of CBF3.

The relatively simple separation methods of gel filtration and velocity sedimentation are valuable tools for studying the structures of proteins and protein complexes. Using these methods, one can derive quantitative information about the native molecular weight, size, and shape of a protein. Interactions between proteins can also be assayed by mixing the proteins together and looking for increases in size or mass. Since hydrodynamics techniques are non-destructive and can be performed using virtually any buffer, native proteins can be analyzed and detected using assays that directly measure protein function.

In this work, I used gel filtration chromatography and velocity gradient sedimentation to characterize the protein components of CBF3 individually and in subcomplexes. I have taken advantage of the finding that recombinant CBF3 proteins produced separately can be mixed together to reconstitute DNA binding activity (Kaplan et al., 1997), indicating that the separately expressed proteins are functional for assembly onto DNA. I report that all of the CBF3 proteins are multimeric in solution and when combined can form subcomplexes, even in the absence of centromeric DNA. I use these results and additional information about size and shape to build a three-dimensional model of CBF3 bound to DNA. I speculate that the subcomplexes formed *in vitro* represent intermediates in the CBF3 assembly pathway, and discuss ways in which the structures of CBF3 components might contribute to DNA binding specificity.

Materials and Methods

Expression of recombinant CBF3 in insect cells

CBF3 genes were subcloned into the pFastBac plasmid (GIBCO BRL, Gaithersburg, MD) and baculovirus DNA was prepared as described in the pFastBac manual. p23, p64, and p110 constructs were created containing the sequence MRGS-H6 fused to the initiator methionine. p110 has methionines at residues 1 and 12 of the published sequence, and I found that p110 fused at Met-12 was more active in bandshifts than protein fused at Met-1. Thus, p110 fused at Met-12 was routinely used for the work described here. Proteins were expressed in High Five insect cells (Invitrogen Corporation, Carlsbad, CA) by infecting with one or more baculoviruses and harvesting after 48 to 72 hours. To minimize protein degradation, all steps of the extraction were performed on ice. Cytoplasmic extracts were prepared by swelling cells in hypotonic

lysis buffer (10 mM Tris-HCl [pH 8.0], 10 mM KCl, 1.5 mM MgCl₂, 10 mM β-mercaptoethanol, 10 μg/ml each of leupeptin, pepstatin, and chymostatin, 50 μM TPCK, 1 mM PMSF), breaking the cells with a dounce tissue grinder (Wheaton, Millville, NJ), and centrifuging the lysate to pellet the nuclei. Supernatants were adjusted to 10% glycerol, 150 mM KCl, and 50 mM β-glycerophosphate before freezing for storage. The nuclear pellet was extracted for 30 minutes in nuclear extraction buffer (10 mM HEPES [pH 8.0], 50 mM β-glycerophosphate, 0.1 mM EDTA, 0.5 M KCl, 5 mM MgCl₂, 10% glycerol, 50 mM NaF, 10 mM β-mercaptoethanol, 10 μg/ml each of leupeptin, pepstatin, and chymostatin, 50 μM TPCK, 1 mM PMSF) and centrifuged to remove DNA and other insoluble material. Under these extraction conditions, p64, p58, and p110 were found primarily in nuclear fractions while p23 partitioned mostly to the cytoplasmic fraction. Nuclear extracts used for purification were prepared in the absence of EDTA so that material could be incubated directly with metal chelate resin (see below). 1% Triton X-100 was included in nuclear extraction buffer when p110 was prepared for purification to improve solubility.

H6-tagged proteins were purified using Ni-NTA Superflow resin (Qiagen, Inc., Valencia, CA). Extracts were adjusted to 3 mM imidazole, added to resin, and bound in batch overnight at 4°. For p64 and p23, resin was washed in Ni buffer (20 mM HEPES [pH 8.0], 500 mM NaCl, 10% glycerol, 10 mM β-mercaptoethanol, 50 mM imidazole), before eluting in the same buffer with 500 mM imidazole. For p110, resin was washed in Ni buffer (3 mM imidazole), eluted in batch with Ni buffer (250 mM imidazole), and diluted to 150 mM NaCl before loading onto a column containing a small volume of

Poros 20 HQ resin (PerSeptive Biosystems, Cambridge, MA) to concentrate the protein. p110 was then eluted in Ni buffer containing 800 mM NaCl and no imidazole.

Gel filtration chromatography

Gel filtration chromatography was performed on a SMART System (Pharmacia Biotech, Uppsala, Sweden) using Superose 6 and Superose 12 columns. To determine diffusion coefficients, standard curves were generated by plotting elution volume versus $1/D$, where D is the diffusion constant, for protein standards (Bio-Rad, Hercules, CA): thyroglobulin (MW = 670 kDa, $D = 2.63 \times 10^{-7} \text{ cm}^2 \text{ s}^{-1}$), bovine gamma globulin (MW = 158, $D = 4.1$), chicken ovalbumin (MW = 44, $D = 7.8$), and equine myoglobin (MW = 17, $D = 11.3$). Diffusion coefficients for protein standards were obtained from Sober (1970) and represent values determined in water at 20°C ($D_{20,w}$). Elution from a gel filtration column correlates with Stokes radius (Siegel and Monty, 1966). The relationship $a = kT/6\pi\eta D$, where a is the Stokes radius; k , the Boltzmann constant; T , absolute temperature; and η , viscosity of the medium, allows $D_{20,w}$ to be determined for an unknown provided that both the unknown and the protein standards are run in the same buffer. All columns were run at 4° with column buffer (10 mM HEPES [pH 8.0], 6 mM MgCl_2 , 10% glycerol, 150 mM NaCl, 10 mM β -mercaptoethanol). Analyses were repeated at least two times and values for $1/D$ varied by no more than 4% between experiments.

Separation of purified proteins

Purified p23 and p110 were diluted to 150 mM NaCl with column buffer lacking NaCl. Because p64 was only partially soluble at 150 mM NaCl, it was diluted to 300 mM NaCl and separated using column buffer containing 300 mM NaCl. Diluted proteins

were centrifuged to remove insoluble material before loading onto the column. Elution was monitored by absorbance at 280 nm and fractions were collected and analyzed by immunoblotting and by bandshift assays.

Separation of proteins in insect cell extract

Because it was difficult to express and purify p58, all analyses of p58-containing mixtures were performed using insect cell extracts. Samples were prepared by dialyzing nuclear extracts into column buffer using a microdialysis system (GIBCO BRL) and then centrifuging to remove insoluble material. To compare directly the elution profiles of different sets of coexpressed proteins, the fraction size, the number of fractions and the starting point for fraction collection were kept constant.

Glycerol gradient sedimentation

Gradients were prepared in 2 ml volumes by layering 0.4 ml of column buffer containing decreasing concentrations of glycerol and incubating the gradient at 4° for 1 hour to equilibrate. Purified p64 was analyzed in 300 mM NaCl and all other proteins were analyzed in 150 mM NaCl. p110 and p64 were sedimented in 15-35% gradients while p23 and p23/p58 were sedimented in 5-25% gradients. Purified p23, p64, and p110 were prepared by diluting to the same salt concentration as the gradient with column buffer lacking glycerol and NaCl. Insect cell extract containing p58 and p23 was dialyzed into column buffer lacking glycerol but containing 150 mM NaCl. For p110 and p64, gradient standards were added directly to the protein samples; for p23 and p58, the standards were run separately and in duplicate. Gradients were centrifuged at 50k rpm 10-16 hrs, 4°, in a TL-S55 swinging bucket rotor (Beckman Instruments, Fullerton, CA) and fractionated by removing 100 µl at a time from the top of the gradient. Fractions

were then analyzed either by bandshifts or by TCA precipitation followed by immunoblotting. Protein standards (Boehringer Mannheim, Indianapolis, IN; cytochrome c [MW = 12.5 kDa, s = 1.9 S], chymotrypsinogen A [MW = 25, s = 2.58], hen egg albumin [MW = 45, s = 3.55], bovine serum albumin [MW = 68, s = 4.22], aldolase [MW = 158, s = 7.4], and catalase [MW = 240, s = 11.3]) were separated on SDS-polyacrylamide gels, stained with Coomassie blue, and quantified using IPLab Gel image processing software (Signal Analytics, Vienna, VA). All samples were analyzed at least two times, and sedimentation coefficients varied by no more than 15% between experiments.

Bandshift assays and immunoblotting

Bandshift assays were performed to determine relative levels of active CBF3 proteins in column and gradient fractions. To detect active p58 in a fraction, an aliquot of that fraction was mixed with excess recombinant p64 and p110 in a standard 30 μ l bandshift reaction containing 40 fmol of DNA probe, 9 μ g casein and 3 μ g sheared salmon sperm DNA in 10 mM HEPES (pH 8.0), 6 mM MgCl₂, 10% glycerol, adjusted to 150 mM KCl. Similarly, to detect p64, fractions were added to excess p23, p58 and p110; p110 was detected in the presence of excess p23, p58 and p64. In all cases, the DNA probe was a 56 base pair fragment of *CEN3* containing CDEIII, 6 nucleotides to the left and 25 nucleotides to the right, synthesized as previously described (Espelin et al., 1997). Samples were separated on 4% nondenaturing acrylamide gels, detected using a phosphorimager, and quantified using ImageQuaNT software (Molecular Dynamics, Sunnyvale, CA).

For immunoblotting, rabbit polyclonal antibodies specific to each of the CBF3 proteins were used and detected using ^{125}I protein A (DuPont/NEN, Boston, MA). Blots were scanned by phosphorimager and quantified as above.

Calculations

Native molecular weights were calculated using the expression $MW = RTs / D(1 - \bar{v}\rho)$, where R is the ideal gas constant; T, absolute temperature; s, sedimentation coefficient; D, diffusion coefficient; \bar{v} , partial specific volume; and ρ , density of water at 20°. Partial specific volumes (0.715 cm^3g^{-1} for p110, 0.724 for p64, 0.697 for p23, and 0.719 for a p23/p58 heterodimer) were estimated from amino acid content. Stokes radii were calculated using the equation $a = kT/6\pi\eta D$, where a = Stokes radius and η = viscosity of water at 20°C. f/f_0 was calculated using the relationship $f/f_0 = a / (3\bar{v}M/4\pi N)^{1/3}$, where M is molecular weight and N is Avagadro's number.

Results

The analysis of recombinant CBF3 proteins expressed in insect cells required experimentally determining two hydrodynamic properties for each protein, the diffusion coefficient and the sedimentation coefficient. I obtained these values using gel filtration columns and glycerol gradients, respectively, each calibrated with protein standards. While other methods may allow more accurate determination of protein hydrodynamic properties, the techniques used here permitted recovery of the samples and subsequent functional analysis. Using the experimentally-determined values, I calculated solution stoichiometries and estimates of both size and shape. A summary of hydrodynamic data for all CBF3 proteins is shown in Table 2.I.

p110 and p64 are homodimers in solution

For the characterization of p110, partially purified protein (Figure 2.1A) was first analyzed by gel filtration chromatography. Absorbance at 280 nm was used to identify the elution volume of the major protein peak, and immunoblotting and bandshift assays confirmed that p110 protein was contained within that peak. An example of a gel filtration experiment is shown in Figure 2.2A. Most of the active p110, identified by bandshift, eluted slightly earlier than the majority of the p110 detected by antibody, suggesting that a portion of the protein was incorrectly modified or improperly folded. p110 eluted just after the 670 kilodalton marker, much larger than expected for an approximately 110 kilodalton monomer. There were two likely explanations for this anomalous behavior. First, p110 might be elongated. The rate of movement through a gel filtration column is determined by the size, or more specifically, by the length, of a protein, which may not correlate directly with molecular weight if the protein is irregularly shaped. Second, if p110 was multimeric in solution, both its size and mass would be greater. To distinguish between these possibilities, I analyzed p110 by glycerol gradient sedimentation (Figure 2.2B). Using the sedimentation coefficient and diffusion coefficient, determined by comparing p110 to protein standards in glycerol gradients and gel filtration chromatography, respectively, I calculated a molecular weight corrected for both size and shape. The corrected, or native, molecular weight of p110 is 197 kilodaltons. By comparing this value to 112 kilodaltons, the monomer molecular weight of p110 predicted from amino acid sequence, I conclude that p110 is a dimer in solution. Based on the diffusion coefficient, a Stokes radius of 80 angstroms was calculated for the p110 dimer. Thus, p110, which directly contacts DNA, is roughly as long as 4.5 helical

turns of a B-form double helix. Finally, by calculating f/f_0 , an indicator of protein shape, I determined that the p110 dimer is elongated, as opposed to globular.

The hydrodynamic properties of p64 were determined using similar methods. Like p110, active p64 ran on a gel filtration column slightly ahead of the peak of total protein (Figure 2.3A). Most of the p64 behaved like a 111 kilodalton protein, although a second, minor, peak detected by bandshift eluted later from the column just ahead of the 44 kilodalton marker. If the primary peak consists of dimeric p64, this secondary peak could be monomeric. I analyzed p64 by glycerol gradient sedimentation (Figure 2.3B) and calculated a native molecular weight of 132 kilodaltons. Thus, the majority of p64 in my preparation is a dimer in solution. p64 has a Stokes radius of approximately 45 angstroms and, unlike p110, is globular.

p23 may be post-translationally modified

Using the same hydrodynamic methods, I analyzed the solution structure of p23. Examples of both gel filtration and glycerol gradient sedimentation experiments are shown in Figure 2.4. Although there was no convenient method for analyzing p23 activity in column fractions, preliminary results demonstrate that when added to an *in vitro* translation reaction of p58, recombinant p23 can activate p58 for CBF3-DNA binding (I. Russell, unpublished results). From the hydrodynamics experiments, the native molecular weight of p23 was determined to be 30.3 kilodaltons. Comparing this value to the predicted monomer molecular weight of 23.9 kilodaltons, a stoichiometry of 1.3 was calculated. While the calculated molecular weight suggests that p23 is a monomer in solution, the value is 30 percent greater than the predicted molecular weight, suggesting that the monomer is heavily post-translationally modified. This notion is

supported by the findings that p23 is homologous to a cytoplasmic glycoprotein in *Dictyostelium* (Kozarov et al., 1995), and that p23 is highly phosphorylated in yeast (K. Kaplan, unpublished observations).

To test the hypothesis that p23 is post-translationally modified in insect cells, I expressed p23 in *E. coli* and purified it for hydrodynamic analysis. I reasoned that because most eukaryotic post-translational modifications do not occur in bacteria, I would see a difference in the molecular weights of p23 expressed in bacteria and p23 expressed in insect cells if the latter protein was post-translationally modified. I applied bacterially-expressed p23 to glycerol gradients and gel filtration columns and calculated a native molecular weight of 25.0 kilodaltons (data not shown, summarized in Table 2.I). Comparing this value to the predicted molecular weight of p23, I conclude that bacterially-expressed protein is a monomer and infer that p23 expressed in insect cells is also a monomer but is post-translationally modified. While the nature of this modification is unknown, p23 in yeast extract migrates on SDS-PAGE closer to p23 from insect cells than it does to p23 from bacteria (data not shown), suggesting that p23 is also modified in yeast. Finally, the presence of a small quantity of p23 eluting earlier from a Superose 12 column than the primary peak (fraction 5, Figure 2.4A) suggests that some p23 may also form a dimer in solution.

p23 and p58 form a heterodimer in solution

I next determined the hydrodynamic properties of p58. Preparation of p58 that is active for CBF3 DNA binding requires coexpression with p23 (Kaplan et al., 1997). Thus, to analyze p58 that was functional, I expressed p58 and p23 together in insect cells and assayed fractions for p58 activity by mixing them with excess p64 and p110 in

bandshift reactions. p58 was very difficult to express and purify, so to have adequate quantities for detection I applied extract prepared from insect cells coexpressing p23 and p58 onto columns and gradients (Figure 2.5) without prior purification. The calculated molecular weight of the material that complemented p64 and p110 in a bandshift was 77.2 kilodaltons. Because p58 that is purified away from p23 is still active for CBF3-DNA binding (Kaplan et al., 1997), it was necessary to determine whether the 77.2 kilodalton active material contained p23 as well as p58.

To determine whether p58 associates with p23 in a p23/p58 coinfection, I took advantage of the observation that p23 fused to glutathione-S-transferase (GST-p23) elutes much earlier from a Superose 12 gel filtration column than does active p58 coexpressed with untagged p23 (data not shown). I reasoned that I could detect an association between p58 and p23 because the peak of active p58 bound to GST-p23 would be shifted to the left relative to the peak of active p58 bound to untagged p23. As shown in Figure 2.5A, when p58 is coexpressed with GST-p23, almost the entire active pool of p58 elutes earlier from the column than the active material in a p23/p58 coinfection. I conclude from this result that in a GST-p23/p58 coexpression most or all of the p58 is physically associated with GST-p23 and I infer that active p58 is also associated with untagged p23 when p58 and p23 are coexpressed. When I add the predicted molecular weight of p58 to the calculated molecular weight of p23 and compare this value to the calculated molecular weight of the p23/p58 complex, I conclude that active p58 is bound to p23 as a heterodimer.

p58 can interact with either p64 or p110

Once I found that I could assay p58 associations in insect cell extract relatively easily using gel filtration chromatography and bandshifts, I decided to apply a similar method to the analysis of interactions between p58 and other CBF3 components. When expressed separately and applied to a Superose 12 column, active p58 in p23/p58 extract peaks in fraction 14 (Figure 2.5A, closed circles), while p64 and p110 peak in fractions 12 and 5, respectively (Figure 2.7A, closed circles, and data not shown). The elution profiles of active p58 coexpressed with either p64 or p110 are shown in Figure 2.6. With extract containing p23, p58, and p64, almost all of the active p58 elutes from the column earlier than it does in a p23/p58 coexpression, peaking in fraction 10 (Figure 2.6A). The completeness of this shift indicates that most of the active p58 in the p23/p58/p64 extract is associated with p64 and suggests that the association is very tight. The coexpression of p23 and p58 with p110 also shifts the elution profile of p58 to the left, demonstrating that p58 and p110 can associate (Figure 2.6B). Two aspects of the p58-p110 association are different from the p23-p58 and p58-p64 associations. First, only a fraction of the p58 is shifted on the p23/p58/p110 column. One explanation for this result might be that p110 is expressed at low levels and therefore is unable to bind up all of the p58. Second, p58 coexpressed with p110 elutes in multiple indistinct peaks spread out over a large portion of the column. This result would be expected if p58 and p110 formed multiple complexes of different sizes. A simpler explanation for both of these observations is that the p58-p110 interaction is unstable, so that p58 that is initially complexed with p110 dissociates during the column run and elutes after the actual p58-p110 complex.

p58, p64, and p110 form a complex in the absence of CEN DNA

I next asked whether I could detect a trimolecular complex that contained p58, p64, and p110. Although it has previously been shown that all three proteins coelute from a gel filtration column (Lechner and Carbon, 1991), those experiments did not demonstrate that the proteins were associated in a single complex. When I compared the fractionation of p23/p58/p110 and p23/p58/p64/p110 extracts on Superose 12 columns, p58 activity profiles were indistinguishable (data not shown). While this result might indicate that a trimolecular complex does not form, I suspected that it was obtained because the resolution of large complexes is limited on a Superose 12 column. As an alternative approach, I asked whether mixing p64 with both p58 and p110 would cause a shift in the elution profile of p64. As shown in Figure 2.7A, p64 in the absence of other CBF3 components elutes as a single peak in fraction 12. Mixing p110 with p64 does not detectably shift the p64 activity profile, indicating that p110 and p64 do not interact in the absence of p58 and p23. When p23/p58 extract is mixed with p64 extract, the majority of the p64 elutes slightly earlier from the column (fraction 10). The peak of p64 activity in this p23/p58 + p64 mix corresponds to the peak of p58 activity in a p23/p58/p64 coexpression (Figure 2.6A), consistent with the formation of a single p58/p64 complex. In contrast to mixing p64 with either p23/p58 or p110 extracts, mixing all three extracts together causes a substantial portion of the p64 to elute very early from the column, in fractions 2-7 (Figure 2.7B). By comparing Figures 2.7A and B, it is apparent that the shifting of p64 into this larger complex is dependent on the presence of both p23/p58 and p110 extracts. This early-eluting p64 has multiple peaks and elutes over a broad volume, suggesting that the interaction that shifts p64 either is relatively weak or consists of

multiple complexes. Because p58 can interact with either p64 or p110, and because I have been unable to detect interactions between p23 and either p64 or p110 (data not shown), I conclude that the shift in p64 activity is due to formation of one or more complexes containing p64, p58, and p110.

p58 in yeast

To examine the relationship between the properties of recombinant CBF3 proteins and the properties of the same proteins in yeast, I applied extract prepared from wild type yeast to a gel filtration column and assayed fractions for p58 activity. In Figure 2.8, I show that the bulk of the active p58 elutes from a Superose 12 column with a peak in fraction 10. This peak of activity corresponds to the peak of active recombinant p58 when it is associated with p64, suggesting that in yeast, the majority of active p58 is associated with p64. Some active p58 in yeast extract elutes earlier from the column, showing that a portion of the p58 in yeast is contained in an even larger complex.

Discussion

CBF3 is an essential DNA binding component of the yeast kinetochore composed of four proteins, but very little is known about how these proteins contribute to the structure of CBF3 and its assembly onto DNA. In this report, I have characterized the CBF3 proteins in solution and shown that p110 is an elongated homodimer, p64 is a globular homodimer, p23 is monomeric, and p23 and p58 form a stable heterodimer. p58 binds to each of the other CBF3 components, and, while there is no evidence that p64 and p110 interact in the absence of p58, p58 can mediate formation of a p58-p64-p110 complex. In conjunction with previous results that describe protein-DNA interactions in the CBF3-DNA complex, I use my characterization of CBF3 protein structures and

protein-protein interactions to build models for both the assembly pathway and the final structure of CBF3 on DNA.

The structure of CBF3

By measuring sedimentation and diffusion coefficients, I have determined the stoichiometries and rough physical dimensions for each of the CBF3 proteins in solution. These descriptions are necessarily imprecise due to a variety of factors, including my reliance on quantitative methods that depend on standard proteins and the limited resolution of the hydrodynamics techniques and detection methods. However, my conclusions about stoichiometry are internally consistent because each of the calculated stoichiometries is very close to integral. Three additional results support the validity of my techniques. First, molecular weights determined for p64 using the two independent hydrodynamics methods are similar (117 versus 138 kilodaltons). Second, the difference in calculated molecular weights of p23 expressed in insect cells and expressed in bacteria is consistent with the difference observed by SDS-PAGE. Third, both p64 and p110 contain coiled-coil sequence motifs, consistent with the finding that these proteins are both homodimers in solution.

Based on the stoichiometries of CBF3 proteins in solution, the simplest proposal for the makeup of the core CBF3-DNA complex is two p64s, one p58, one p23, and two p110s. The predicted total molecular weight of the protein components of this complex is 405 kilodaltons. A model for the assembled structure based on stoichiometries, Stokes radii, and protein-protein interactions, as well as protein-DNA interactions previously identified by crosslinking studies (Espelin et al., 1997), is shown in Figure 2.9. I assumed uniform protein density and circular cross-sections to calculate the shorter

dimensions of the proteins in this model. This assumption is an oversimplification, but it enables an approximation of what the complex might look like. The CBF3-DNA complex is 180 angstroms long, with a width ranging from 55 to 120 angstroms. For comparison, a single nucleosome is 110 angstroms in diameter and 40 angstroms thick. My results do not rule out that the three-dimensional structure of CBF3-DNA is more complicated and might include bending or twisting of the DNA and proteins.

Assembly of CBF3

My results demonstrate that p58 plays a central role in the assembly of CBF3 by interacting with each of the other CBF3 proteins and mediating interactions among them. p58 forms stable complexes with both p23 and p64. A tight interaction with p23 has previously been suggested by findings that p23 and p58 copurify when expressed in yeast and when expressed in insect cells (Stemmann and Lechner, 1996; Kaplan et al., 1997). Based on my findings that all active p58 is associated with p23 when the proteins are coexpressed in insect cells, and that all CBF3-DNA complex formed in the presence of p23 contains p23 (data not shown), I infer that all of the partial complexes that include p58 also include p23. It has been shown that p23 can be purified away from activated p58, leaving the p58 competent to assemble into CBF3 in the absence of p23 (Kaplan et al., 1997). This result demonstrates that while p23 and p58 can interact, a stoichiometric interaction is not required for CBF3 assembly. In other words, p23 plays a catalytic role, but not a structural role, in CBF3. Why, then, is p23 included in the CBF3-DNA complex if it isn't required? Presumably, p23 binds to p58 as a necessary step in the post-translational activation of p58. That this interaction is required for CBF3 assembly ensures that p23 and p58 will interact, and may lead to the relocalization of p23 to the

kinetochore. I speculate that this localization event enables p23 to perform an additional function at the kinetochore possibly related to its role in the SCF ubiquitin ligase complex (Feldman, et al., 1997; Skowyra, et al., 1997). Mutations in p23 have been identified that separate its function at the G1-S transition, where it is required for the degradation of Sic1p, and at G2-M, where p58 must be activated for kinetochore assembly (Connelly and Hieter, 1996). It will be interesting to determine whether mutations in p23 can be found that arrest cells at G2-M but still enable p23 to activate p58.

A stable interaction between p58 and p64 has been suggested by the finding that p58 copurifies with p64 from yeast extract (Stemmann and Lechner, 1996). This notion is supported by two of my results. First, all of the active p58 associates with p64 when the proteins are expressed in insect cells. Second, the active p58 in yeast extract is present in a complex that behaves in gel filtration chromatography similarly to the p58-p64 complex in insect cells. In yeast, p58 is turned over very rapidly (Kaplan et al., 1997). Preliminary results showing that the efficient overexpression of p58 *in vivo* requires the co-overexpression of p64 (A. Grancell, unpublished observation) suggest that the tight interaction between p58 and p64 serves to stabilize p58 against degradation. Such an interaction might facilitate the folding of p58 and prevent its degradation by machinery that scavenges misfolded proteins.

Based on my observations of interactions among CBF3 proteins, I propose that the various p58-containing subcomplexes represent intermediates in the CBF3 assembly pathway (Figure 2.10B). A p23-p58-p64 complex appears to be the dominant form of active p58 in the cell. Because protein complexes containing p110 appear to be unstable, I suspect that interactions between the p23-p58-p64 complex and p110 are transient, until

the four proteins together contact DNA. The DNA-bound CBF3 is stabilized by protein-DNA interactions, by protein-protein interactions, and possibly by conformational changes induced in the proteins or DNA upon CBF3 binding. A prediction of this model that can be tested is that binding to DNA will serve to stabilize protein-protein interactions in a complex containing p58, p64 and p110.

Implications for the specificity of CBF3 binding

In a haploid yeast cell, there are only 32 sites where CBF3 must assemble (centromeres). Given the large amount of additional DNA in the genome, what characteristics of CBF3 determine its binding specificity? In one model, CDEIII contains essential nucleotides that interact with each of the CBF3 proteins, so that assembly of CBF3 onto DNA involves specific interactions both among CBF3 proteins and between each of the proteins and DNA. Thus, CBF3 is like a puzzle of interlocking pieces. Two predictions of this model are 1) nucleotides that contact each of the proteins will be highly conserved and 2) mutating any of the bases involved in specific contacts would disrupt CBF3 assembly. An alternative model for CBF3 binding specificity is that only one of the CBF3 proteins makes essential sequence-specific contacts, and that the other proteins act to facilitate this binding. Predictions of this model include 1) the nucleotides contacted by the sequence-specific binding protein are highly conserved while other nucleotides are less well-conserved, and 2) mutating the contact sites for the sequence-specific binding protein will have greater impact on CBF3 binding than mutating bases that contact other proteins.

The second model is more strongly supported by both *in vivo* mutational analysis of CDEIII and protein-DNA crosslinking results. p64 contacts four of the seven

absolutely conserved nucleotides in CDEIII, and two more completely conserved bases are immediately adjacent to essential p64 contacts (Espelin et al., 1997). Additionally, point mutations in the bases contacted by p64 have more severe consequences for *in vivo* function than do mutations in bases contacted by either p58 or p110 (Jehn et al., 1991). These results suggest that p64 may be primarily responsible for determining the binding specificity of CBF3.

My finding that p64 is a dimer in solution also supports the second hypothesis. I propose that p64, like other zinc cluster proteins, specifically recognizes DNA sequences at its binding site, and by this functions as the sequence recognition component of CBF3. Other members of the zinc cluster family bind as homodimers to sequences containing either direct or inverted repeats of CCG (Marmorstein et al., 1992; Marmorstein and Harrison, 1994; Zhang and Guarente, 1994). The requirement for intact CCGs is not absolute. *In vitro* selection of binding sites was performed for the zinc cluster protein Hap1p (Zhang and Guarente, 1994). Many of the selected sequences contained only one intact CCG, and the only base conserved in the second would-be CCG was the central nucleotide. CDEIII sequences contain a pattern similar to the patterns of isolated Hap1p binding sites. One CCG, which is a contact site for p64, is absolutely conserved, while the other p64 contact site contains a single completely conserved G near the center of the site (see Chapter 1). These p64 contact regions form a pseudo-palindromic recognition site centered on the conserved G residue contacted by p58. I propose that, like Hap1p, p64 binds as a dimer to the intact CCG and to the imperfect G-containing half site. Furthermore, I propose that each of these interactions is mediated by zinc clusters. Because Hap1p binds to its imperfect recognition site *in vitro* in the absence of other

protein factors, I speculate that the reason p64 does not bind to CDEIII on its own is not because the recognition site is imperfect, but because in the absence of other CBF3 proteins p64 is in an incorrect conformation for DNA binding. Thus, interacting with p58 and p110 leads to a conformational change in p64 that enables p64 to bind to CDEIII. This obligatory interaction causes p64 to “recruit” p58 and p110 to the centromere, where they may function to stabilize CBF3 by binding non-specifically to DNA or as assembly factors for other components of the kinetochore. This model is presented in schematic form in Figure 2.10. A prediction of this model is that formation of a p58-p64-p110 complex changes the conformation of p64.

Future directions

In addition to understanding the structure of CBF3, I would like to know how CBF3 assembly is regulated and how CBF3 contributes to assembly of the rest of the kinetochore. Kaplan, et al. (1997) have shown that p58 is regulated by a combination of activation and destruction mechanisms. In light of the results presented in this chapter demonstrating that p58 plays a central role in CBF3 assembly, the finding that p58 is highly regulated suggests that controlling p58 activity effectively controls assembly of the entire CBF3 complex. Now that I have dissected the interactions made by p58, it will be interesting to determine which step in the CBF3 assembly pathway, if any, is effected by p58 activation. This assembly of CBF3 onto DNA appears to be an initiating event for kinetochore assembly. The structural features of the CBF3 complex suggest that p64 and p110 are the most accessible proteins for interaction with other kinetochore components. Using this information, I might now target p64 and p110 for mutational

analysis to determine how these proteins contribute to interaction with other kinetochore proteins and with centromeric DNA sequences outside of CDEIII.

Acknowledgements

I thank Ken Kaplan for donating p23 purified from insect cells and I am grateful to members of Tania Baker's laboratory for their help with the SMART System.

Table 2.I. Summary of protein hydrodynamics results.

	D (x 10 ⁻⁷ cm)	s (Svedbergs)	calculated MW (kilodaltons)	monomer MW* (kilodaltons)	stoichiometry‡	Stokes radius (angstroms)	f/f ₀ ‡‡
p110	2.66 (2.64-2.68)	6.14 (5.70-6.58)	197.2 (184.4-209.7)	112.2	1.76 (1.64-1.87)	80.1 (79.5-80.7)	2.09 (2.03-2.15)
p64	4.75 (4.69-4.80)	7.09 (6.70-7.49)	131.7 (126.0-137.6)	73.9	1.78 (1.71-1.86)	44.8 (44.4-45.4)	1.34 (1.30-1.37)
p58/p23	6.00 (5.70-6.34)	5.26 (4.95-5.58)	76.0 (75.2-76.3)	56.3 [86.6***]	0.88*** (0.87-0.88)***	35.5 (33.6-37.4)	1.27 (1.20-1.34)
p23	7.02 (6.90-7.15)	2.65 (2.49-2.81)	30.3 (29.0-31.6)	23.9	1.27 (1.21-1.32)	30.3 (29.8-30.9)	1.48 (1.43-1.52)
	7.01** (6.89-7.14)**	2.18** (2.02-2.33)**	25.0** (23.6-26.2)**	23.9	1.05** (0.99-1.10)**	30.4** (29.8-30.9)**	1.58** (1.52-1.63)**

Numbers in parenthesis indicate the range for each value as dictated by the size of gel filtration or glycerol gradient fractions.

* Calculated from the predicted amino acid sequence of protein plus 6H tag, if used.

** Indicates values for p23 expressed in bacteria.

*** Indicates values that include the monomer molecular weight of p58 plus the calculated molecular weight of p23.

‡ Determined by dividing calculated MW by monomer MW.

‡‡ Proteins with f/f₀ ratios less than 1.5 are considered “globular”, while proteins with ratios greater than 1.5 are considered “elongated”.

Figure 2.1. Recombinant CBF3 components. (A-C) Recombinant H6-tagged p110 (A), p64 (B), and p23 (C) purified from baculovirus-infected insect cells were separated on SDS-polyacrylamide gels and stained with Coomassie blue. Arrows indicate the full-length proteins. The p110 preparation contains a contaminating protein migrating at ~48 kilodaltons which is detectable by anti-p110 immunoblotting, suggesting that it is a proteolytic breakdown product of p110. The proteins shown here were used for quantitative hydrodynamic analysis. (D) Recombinant p58 used for quantitative hydrodynamics was undetectable by Western blot and was assayed by bandshift. Extract from baculovirus-infected insect cells expressing p58 and p23 was mixed with p64, p110, and CDEIII-containing radiolabelled DNA probe. Lane 1, no p23/p58 extract added; lane 2, p23/p58 extract added.

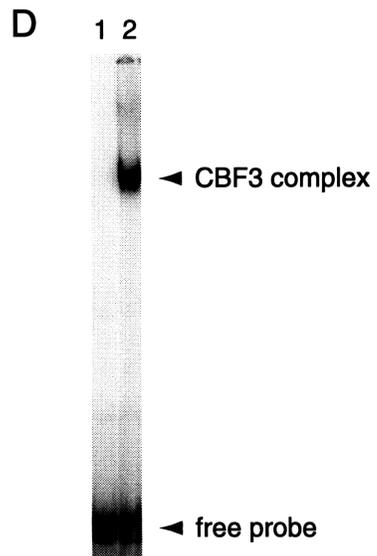
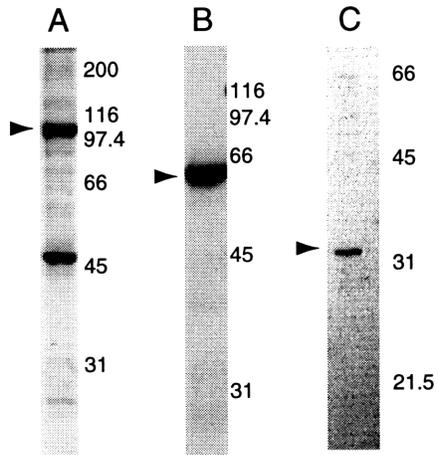
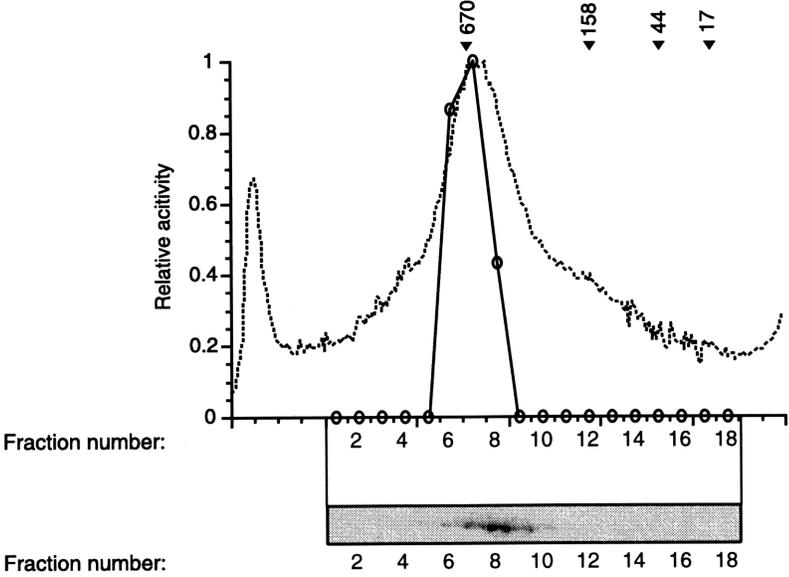


Figure 2.2. Hydrodynamic analysis of p110. (A) The A_{280} elution profile (dotted line) of purified p110 applied to a Superose 6 gel filtration column is shown. The solid line with circles indicates relative *in vitro* CDEIII DNA binding activity of each fraction when mixed with excess p23, p58, and p64. Below the graph, the fractions are analyzed by immunoblotting with affinity-purified anti-p110 antibody and ^{125}I protein A. Molecular weights of protein standards are indicated. (B) Purified p110 was sedimented in a 15-35% glycerol gradient and protein from each fraction was TCA precipitated, separated on an SDS-polyacrylamide gel, and immunoblotted as in A. Molecular weights of protein standards are indicated.

A Superose 6 gel filtration



B Glycerol gradient sedimentation

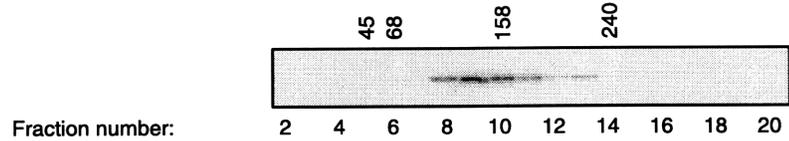
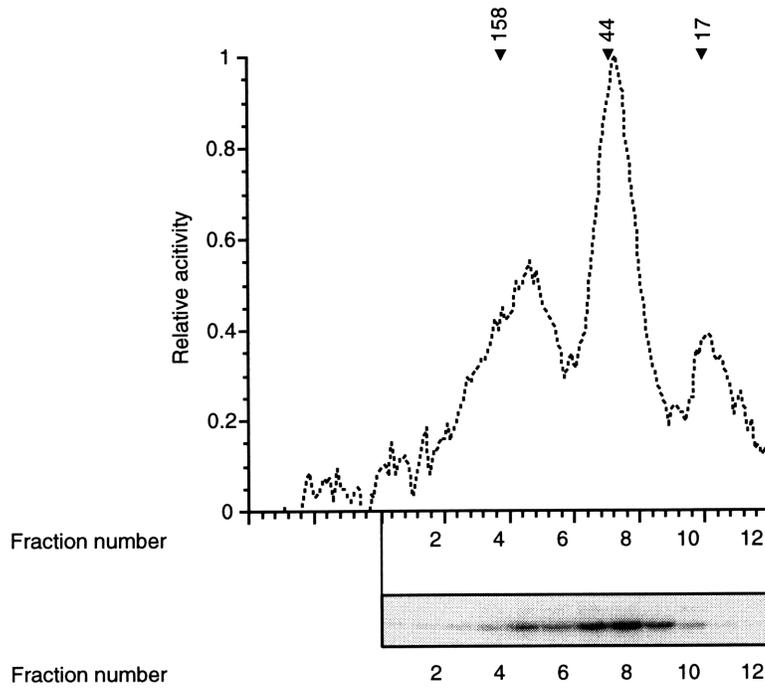


Figure 2.3. Hydrodynamic analysis of p64. (A) Purified p64 was applied to a Superose 12 gel filtration column. The A_{280} elution profile (dotted line) and relative *in vitro* DNA binding activity of fractions mixed with excess p110, p23, and p58 (solid line with circles) are shown. Below the graph, fractions were immunoblotted with anti-p64 antibody and detected with ^{125}I protein A. Molecular weights of protein standards are indicated. (B) Purified p64 was sedimented in a 15-35% glycerol gradient. TCA-precipitated protein from each fraction was separated by SDS-PAGE and immunoblotted as in A. Molecular weights of protein standards are indicated.

A Superose 12 gel filtration



B Glycerol gradient sedimentation

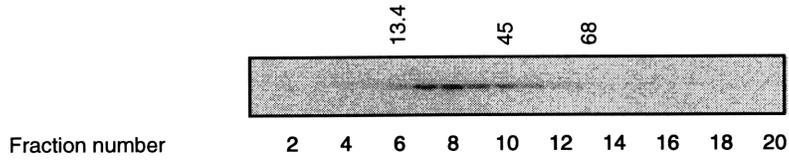
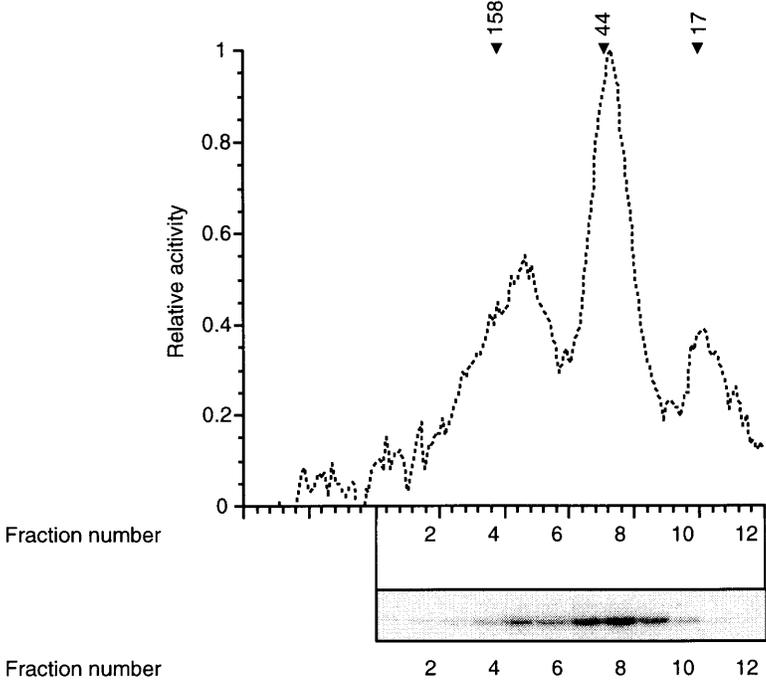


Figure 2.4. Hydrodynamic analysis of p23. (A) Purified p23 was applied to a Superose 12 gel filtration column and the A_{280} elution profile (dotted line) is shown. Fractions were also analyzed by immunoblotting using anti-p23 antibodies and detected with ^{125}I protein A. Molecular weights of protein standards are indicated. (B) Purified p23 was sedimented on a 5-25% glycerol gradient. Fractions were TCA-precipitated and analyzed by immunoblot as in A. Molecular weights of protein standards are indicated.

A Superose 12 gel filtration



B Glycerol gradient sedimentation

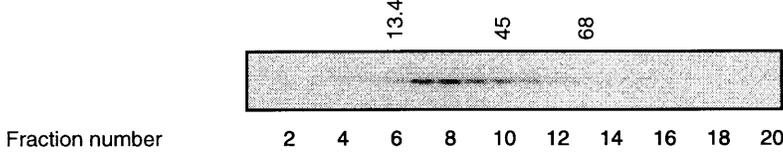
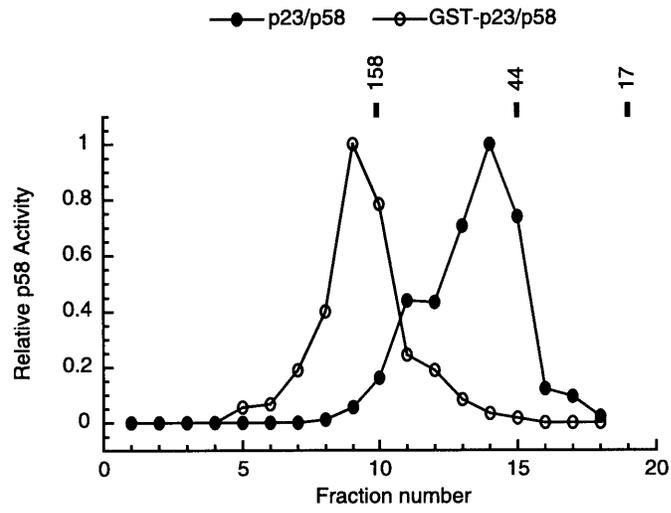


Figure 2.5. Hydrodynamic analysis of p58/p23. (A) Superose 12 gel filtration chromatography of p58 coexpressed with different forms of p23. Extract from baculovirus-infected insect cells expressing p58 and either p23 (closed circles) or GST-p23 (open circles) was applied to a Superose 12 gel filtration column. The amount of active p58 was assayed by mixing a portion of each fraction with excess p64 and p110 in a CDEIII bandshift assay, and the amount of probe shifted was quantified. Molecular weights of protein standards are indicated. (B) Extract from baculovirus-infected insect cells coexpressing p23 and p58 was sedimented in a 5-25% glycerol gradient. Fractions were assayed for active p58 as in part A. Molecular weights of protein standards are indicated.

A Superose 12 gel filtration



B Glycerol gradient sedimentation

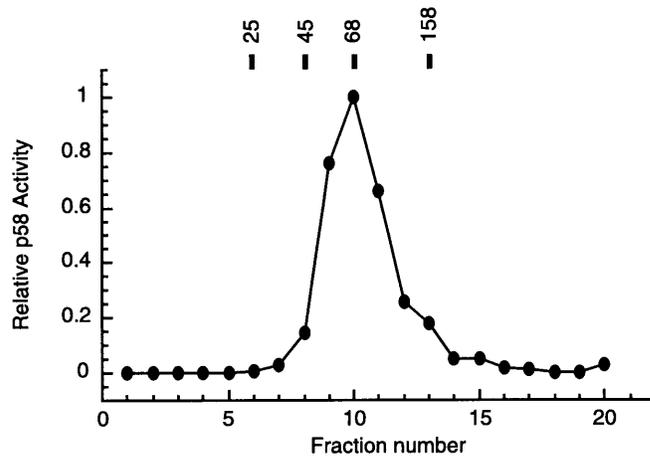
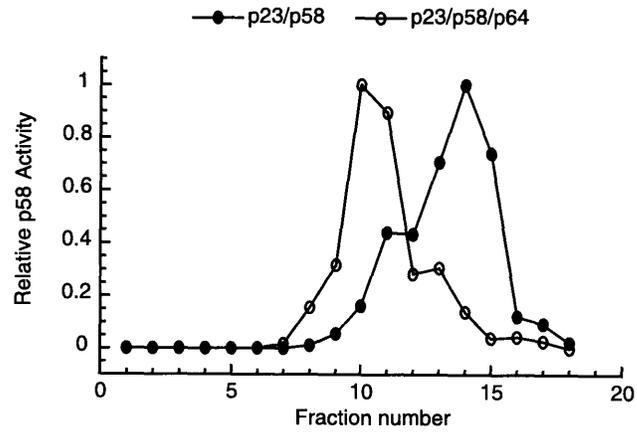


Figure 2.6. Coexpression with p110 or p64 increases the apparent size of p58. Insect cell extracts containing p23, p58, and either p64 or p110 were applied to Superose 12 gel filtration columns and assayed for p58 activity as described in Figure 2.5. For comparison, the p58 activity profile of p23/p58 extract is also shown (closed circles). (A) Extract containing p23, p58, and p64 (open circles). (B) Extract containing p23, p58, and p110 (open circles).

A



B

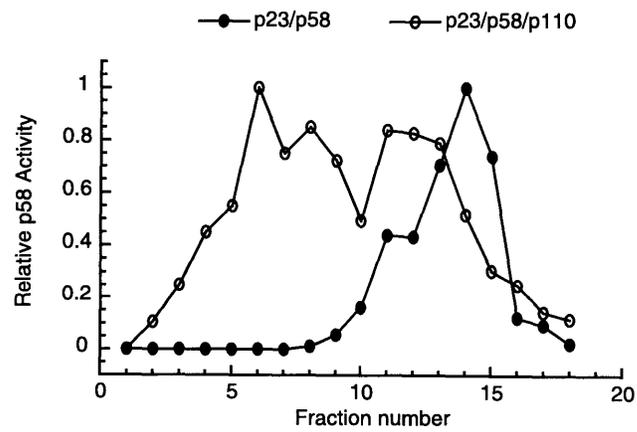
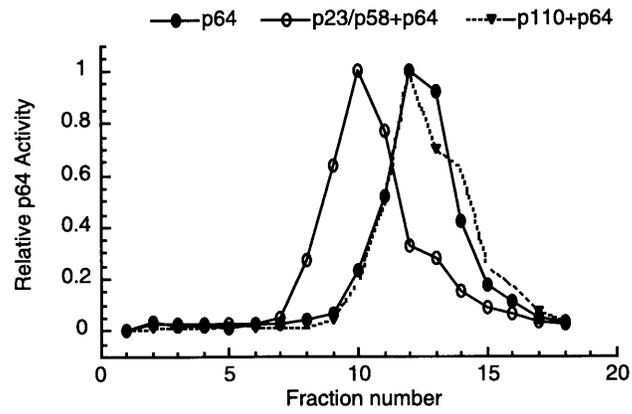


Figure 2.7. Mixing with p23, p58, and p110 but not with p23 and p58 or p110 increases the apparent size of p64. Insect cell extract containing p64 was mixed with extract containing p23 and p58, with extract containing p110, or with both extracts, separated on a Superose 12 gel filtration column and assayed for p64 activity by mixing fractions with excess p23, p58, and p110 in a CDEIII DNA bandshift. The amount of probe shifted was quantified and graphed. (A) p64 mixed with p23/p58 (open circles) elutes slightly earlier from the column than does p64 alone. The elution profile of p64 does not change in the presence of p110 (triangles). (B) Mixing p64 extract with p23/p58- and p110-containing extracts (open circle) causes some of the p64 to elute much earlier from the column.

A



B

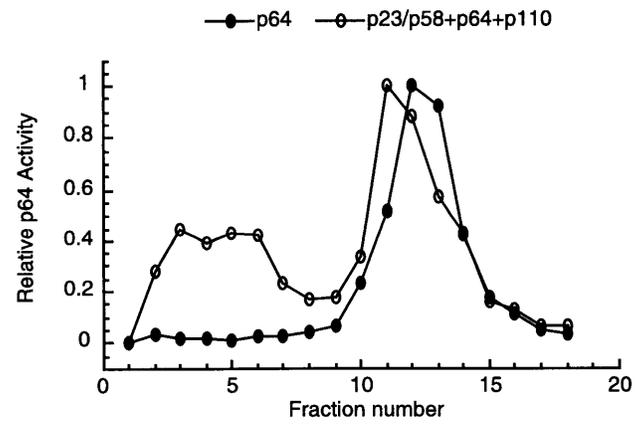


Figure 2.8. p58 in yeast elutes earlier from a gel filtration column than recombinant p58. Wild type yeast extract (triangles), or extract from insect cells expressing p23 and p58 (open circles) or p23, p58, and p64 (closed circles) was applied to a Superose 12 gel filtration column and fractions were assayed for active p58 as described in Figure 2.5.

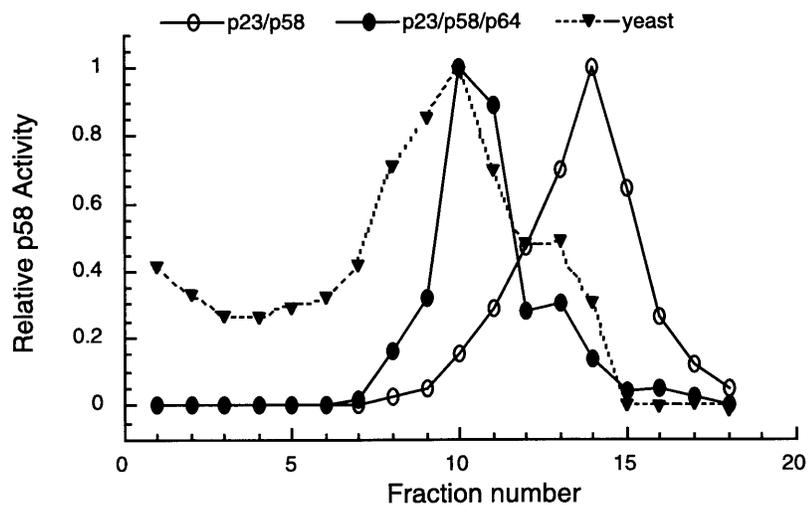
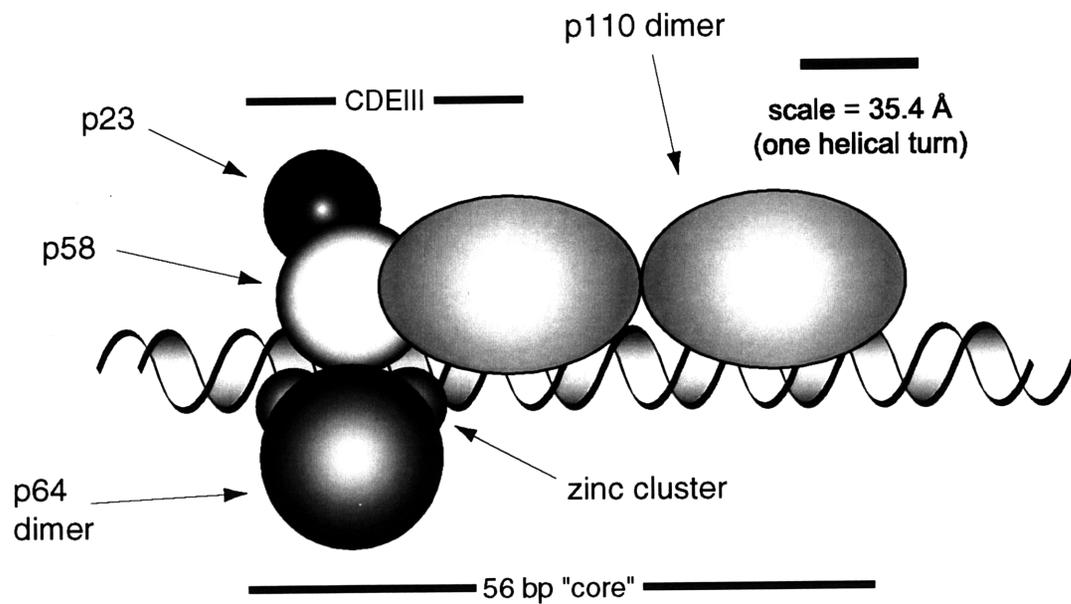
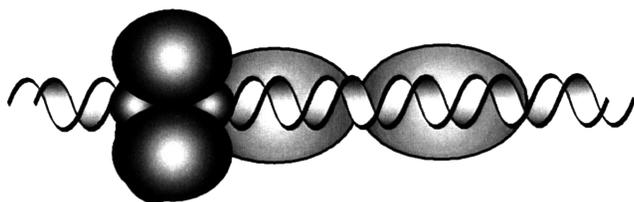


Figure 2.9. A scale model for the structure of CBF3 bound to DNA. CBF3 proteins with experimentally-determined lengths and stoichiometries are positioned on CEN DNA according to protein-DNA crosslinking results (Espelin et al., 1997). Proteins are drawn to scale based on calculated Stokes radii, assuming constant protein density and circular cross-sections. The 56 base pair "core" DNA is the region protected from nuclease digestion by CBF3 *in vitro*.



Bottom-up view



Top-down view

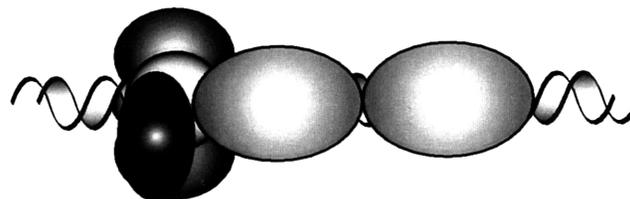
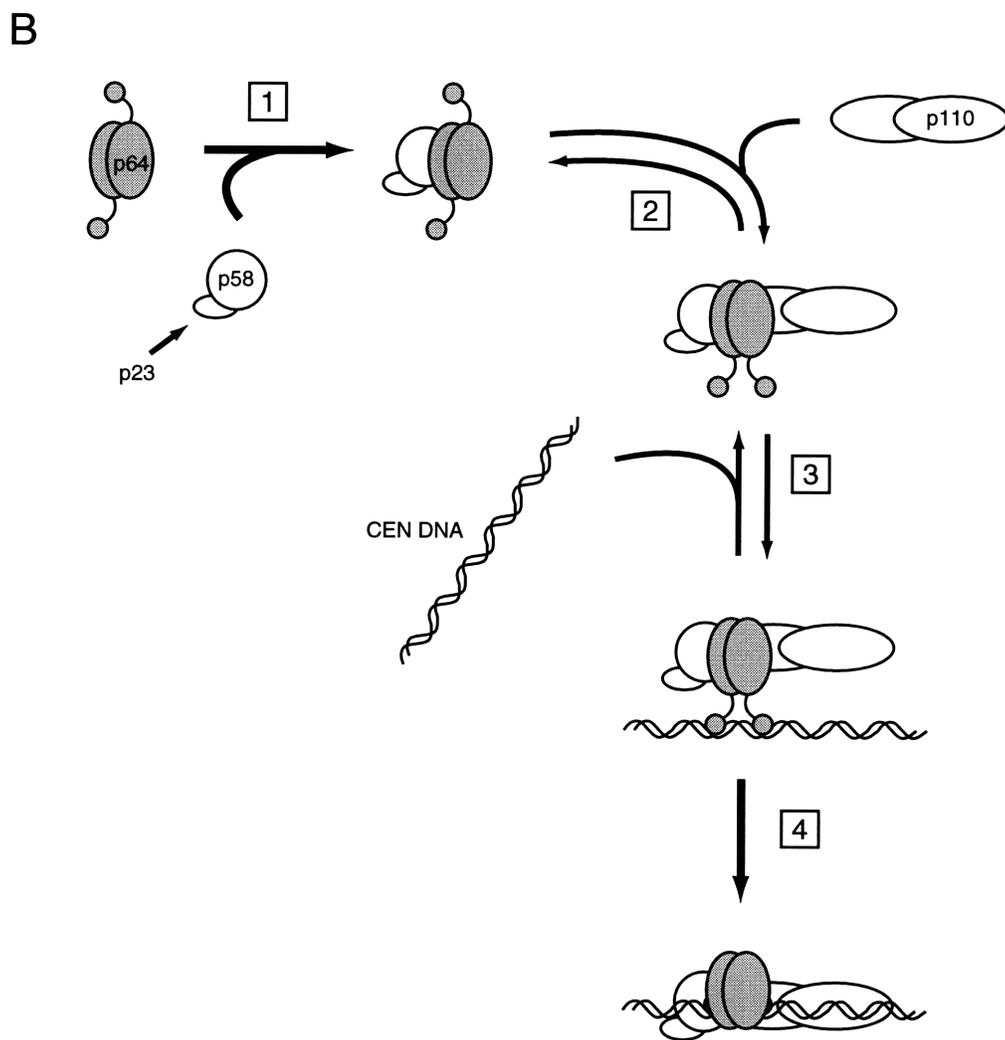
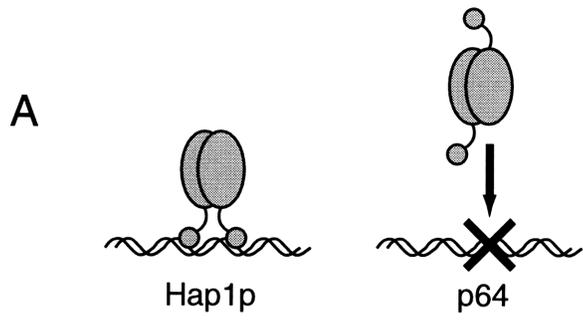


Figure 2.10. A speculative model for the role of p64 in CBF3 assembly onto DNA.

(A) Most zinc cluster proteins, including Hap1p, bind DNA as heterodimers in the absence of other protein factors. p64 cannot bind CDEIII without the other CBF3 proteins. I speculate that the reason p64 cannot bind DNA alone is because it is not in the correct conformation. (B) p64 associates first with p58 and p23 to form a stable complex (1). This complex is still unable to bind DNA, but it can transiently interact with p110 (2). Formation of this tetramolecular complex leads to a conformational change in p64, which enables p64 to bind centromeric DNA (3). Binding of p64 to DNA alters the conformation of all CBF3 proteins and possibly the DNA, locking the nucleoprotein complex together (4).



Chapter 3

Final thoughts: A model for kinetochore assembly

In the previous chapter, I described experiments probing the structure of CBF3 *in vitro*. These experiments led me to a specific model for the assembly of CBF3 in which p58 mediates associations between the other CBF3 proteins and p64 provides the specificity for DNA binding by recognizing specific regions of CDEIII. However, while CBF3 is clearly an essential component of the yeast kinetochore and may even function to initiate kinetochore assembly, it is not sufficient to mediate the binding of a chromosome to the mitotic spindle. In this chapter, I present a model for the assembly of the kinetochore based on these observations and on published information about demonstrated and suspected components of the kinetochore.

A schematic representation of this model is presented in Figure 3.1. CBF3 functions as an initiator complex by selectively interacting with CDEIII DNA, similar to the way in which ORC initiates assembly of replication machinery at the replication origin (Bell and Stillman, 1992), and TFIID initiates assembly of transcriptional proteins at promoters (reviewed in Orphanides et al., 1996). CBF3 binds to CDEIII *in vivo* in the absence of other centromeric sequences and presumably without requiring other kinetochore proteins. Once bound, CBF3 recruits additional kinetochore components, including microtubule-binding proteins and Mif2p.

The next step in kinetochore assembly is a conformational change of the centromeric DNA facilitated by CBF3 and the proteins that initially bind to CBF3. Ample evidence suggests that changes in chromatin structure are required for the assembly and maintenance of the kinetochore, including the findings that genetic manipulations that alter chromatin structure can increase chromosome loss rates and that Cbf1p may be involved in remodeling chromatin at sites of transcription. I propose that Mif2p, which has strong genetic interactions with Cbf1p, initiates bending in CDEII that is enhanced by recruitment of additional molecules of Mif2p and Cbf1p. This bending leads to formation of a loop that encompasses the 150-250 base pairs of DNA protected from nuclease digestion by the kinetochore. A loop of 200 base pairs of B-form DNA would have a circumference of approximately 65 nm. This size is important to support end binding of a microtubule (see below).

Formation of the loop complex allows the assembly of additional structural and microtubule-binding components onto the kinetochore. The entire centromere (CDEI-CDEIII) supports more efficient attachment to microtubules *in vitro* than does CDEIII alone. One interpretation of this result is that proteins that bind CDEIII form a single unit of microtubule attachment and that this unit is repeated multiple times when CDEIII is in the context of other centromere sequences. p110 is a strong candidate for a factor that polymerizes along the remaining centromeric DNA and mediates the binding of microtubule-attachment units, because additional p110 molecules appear to polymerize onto CBF3 and bind non-specific DNA in the CBF3 “extended” complex.

While kinetochore proteins can mediate binding of a small piece of CEN DNA to microtubules *in vitro*, it is likely that the kinetochore requires stabilization by association

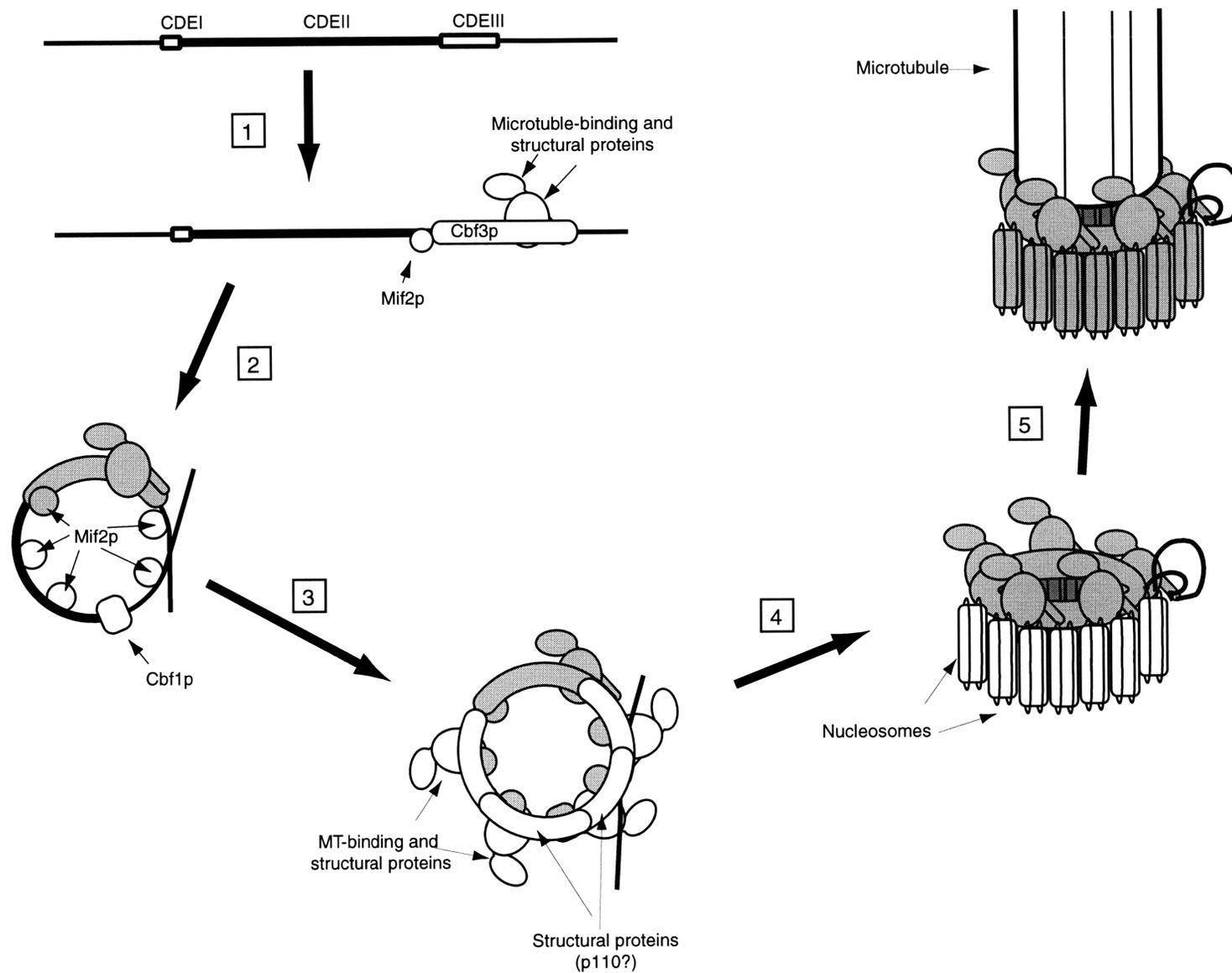
with additional DNA and DNA-bound proteins *in vivo*. The kinetochore transduces forces generated by the microtubule, or at the site of kinetochore-microtubule interaction, to the chromosome. Thus, the kinetochore has stricter stability requirements than most other nucleoprotein complexes. I propose that additional stabilization of the kinetochore is achieved by binding of nucleosomes to the kinetochore, in a structure that may resemble the chromatin “30 nanometer fiber” of organized nucleosomes. These adjacent nucleosomes may contain Cse4p, a histone H3 homolog that is similar to the mammalian centromere protein CENP-A. In mammalian cells, the kinetochore also both contains centromeric DNA and is associated with surrounding specialized chromatin (heterochromatin).

Once the kinetochore is fully assembled it can interact with a microtubule. In mammalian cells, microtubules attach to the kinetochore by inserting into the fibrous corona surrounding the kinetochore outer plate (see Appendix B). It has been proposed that long, flexible proteins, including the motor CENP-E, surround the microtubule and extend part way up it to make contacts along the microtubule wall (Yao et al., 1997). I suggest that a similar interaction occurs at the yeast kinetochore, which has also been shown to form an end-on attachment with the microtubule. The kinetochore loop structure proposed here is slightly larger in circumference than a microtubule (65 nm versus 63 nm), so that microtubule-binding proteins extending from the loop could surround the microtubule end.

To test this model, I would ask several specific questions about the structure of the kinetochore. First, what are the identities of the kinetochore components that bind to microtubules? Once identified, it should be possible to use *in vitro* microtubule binding

assays to determine the order of addition of these components. Second, what is the chromatin structure at the kinetochore, and what components of the kinetochore are required to define this structure? Specific chromatin alterations might include bending, looping, and melting. Finally, what is the structure of the kinetochore as a whole? The overall shape and size of kinetochores formed *in vivo* and excised using engineered restriction sites would be analyzed using hydrodynamic methods or scanning electron microscopy.

Figure 3.1. Speculative model for assembly of the kinetochore. See text for description.



Appendix A

Physical domain mapping of p64 and p110 by partial proteolysis

Introduction

How does a large nucleoprotein structure assemble? CBF3 contains at least three different polypeptides in varying stoichiometries and contacts a large region of DNA. All of the protein components are required for the assembly of CBF3 onto DNA (Kaplan et al., 1997) and most point mutations in CDEIII impair centromere function *in vivo* (Jehn et al., 1991). These results demonstrate that multiple protein-protein and protein-DNA interactions are required to correctly assemble the CBF3-DNA complex. Understanding how each of these interactions contributes to the function of CBF3 at the kinetochore will require determining not only which proteins associate with which other proteins, or which proteins contact what region of DNA, but also what parts of the proteins mediate these interactions.

One approach to determining the relationship between protein structure and function relies on the observation that many proteins contain functionally distinct physical domains. That is, to some extent, proteins are modular. This point has been demonstrated convincingly for transcription factors that possess separate DNA-binding and transcriptional activation domains, which can be exchanged to create hybrid functional proteins with novel specificities (Reece and Ptashne, 1993). It is not

surprising that these domains are often composed of contiguous sequences of amino acids that can fold independently.

Using the notion that protein functional regions constitute separable physical domains, protein domain organization has been studied using limited proteolysis (Nakayama et al., 1987). Flexible regions of proteins that link folded structural domains are more accessible to endoproteases than sequences that compose structural units. Thus, by treating proteins with frequently-cutting proteases it is possible to map protein structure by delineating physical domains. With the addition of N-terminal microsequencing, the exact boundaries of the domains may be identified.

Here, I explore the domain organization of the CBF3 proteins p64 and p110. I find that p64 contains at least two distinct physical regions, one of them being the zinc cluster domain that putatively binds the conserved CCG sequence of CDEIII. p110 contains three domains, that may constitute a dimerization domain, a DNA-binding domain, and a domain that interacts with the other CBF3 components.

Materials and Methods

Expression and purification of recombinant p64 in *E. coli*

E. coli strain BL21 containing the plasmid pUBS520, which expresses the DnaY tRNA-Arg_{AGG/AGG} gene (Brinkmann et al., 1989), was transformed with *CEP3* cloned into the vector pQE30 (Qiagen). Expression of the resulting MRGS-H6-tagged p64 was induced with IPTG at 30°. Cells were washed and sonicated in 50 mM NaPO₄, 300 mM NaCl before extract was bound to Ni-NTA Superflow resin. Remaining purification steps were performed as described in Chapter 2.

Purification of p110

p110 purified as described in Chapter 2 was insufficiently pure for proteolytic analysis. Partially purified p110 was diluted to 150 mM NaCl and applied to a Poros 20 HQ column using a BioCAD chromatography system (PerSeptive Biosystems). After washing, protein was eluted in a 150 mM to 1 M NaCl gradient. Fractions containing p110 were identified by SDS-PAGE and silver staining.

Partial proteolysis

p64 and p110 were treated with dilute proteolytic enzymes in 50 mM HEPES (pH 8.0), 300 mM NaCl, 10% glycerol, 1 mM EDTA. Initial tests with different concentrations of trypsin, chymotrypsin, or *S. aureus* V8 protease (Worthington Biochemicals, Freehold, NJ) were performed to determine an approximate concentration for generating proteolytic fragments. To start reactions, enzyme was added and reactions incubated at 30°. Aliquots were removed at different time points and proteolysis was stopped by boiling the sample in SDS-PAGE loading buffer and 5 mM PMSF. After separating samples on SDS-polyacrylamide gels, protein was detected by Coomassie staining (p64) or silver staining (p110). Microsequencing of p64 proteolytic fragments was performed at the MIT Biopolymers Lab (Cambridge, MA).

Results

p64 expressed in *E. coli* and in insect cells behaves the same in *in vitro* DNA binding

p64 was expressed in *E. coli* as a H6- fusion protein and purified. I compared the bandshift activity of H6-p64 expressed in bacteria to the same protein expressed in insect cells and found that the two proteins have approximately the same specific activity when

complementing p58, p23, and p110 in a bandshift reaction (Figure A.1). This result demonstrates that p64 expressed in bacteria is active and therefore structurally intact.

The p64 zinc cluster is separable from the rest of p64

Purified p64 was treated with trypsin, chymotrypsin, and staphylococcus V8 protease for varying lengths of time (Figure A.2). By primary sequence analysis, p64 contains 70, 74, and 78 possible cleavage sites, respectively, for these proteases distributed throughout the 608 amino acid sequence. All three enzymes gave a similar primary digestion product that migrated at approximately 58 kilodaltons (Figure A.2A). N-terminal sequence analysis of this fragment showed that the cleavage site was slightly downstream of the zinc cluster domain, as shown in Figure A.2B. No fragment corresponding to the cleaved zinc cluster was detected, most likely because the fragment would have migrated off the bottom of the gel. The finding that multiple enzymes produce similarly-sized products supports the notion that the sites of cleavage are exposed regions that link structural domains.

Given this result, I constructed an H6-tagged version of p64 that lacked regions upstream of the proteolytic cleavage site. This protein expressed successfully in insect cells but failed to function in a bandshift assay when mixed with p23, p58, and p110 (data not shown). This result was expected, since the zinc clusters of other family members are required for DNA-binding (Keegan et al., 1986; Pfeifer et al., 1989), and mutations in highly conserved residues in the zinc cluster of p64 lead to increased chromosome loss *in vivo* (Lechner, 1994). The next step in analysis of p64 structure will be to identify other structural domains in the remaining 58 kilodaltons of the protein.

p110 contains three physical domains

I applied the same method of partial proteolysis to p110 purified from insect cells (Figure A.3A). Full length p110 behaves on SDS-PAGE as a 116 kilodalton protein. Treatment with dilute trypsin generates fragments of 96, 72, 40, and 25 kilodaltons. Based on the kinetics of formation of the proteolytic products, a model for the domain organization of p110 can be proposed (Figure A.3B). During the period of proteolytic digestion, the strongest band that develops has a size of 72 kilodaltons. Formation of this product corresponds with the appearance of the 40 kilodalton fragment. I reasoned that because 40 kilodaltons plus 72 kilodaltons is approximately 116 kilodaltons, these fragments are probably formed directly by cleavage of the full length protein. A decrease in the amount of the 96 kilodalton fragment corresponds with an increase in the amounts of the 25 kilodalton and 72 kilodalton fragments, suggesting that these two components constitute the 96 kilodalton product. I failed to detect an expected 20 kilodalton fragment, most likely because it would have migrated off the gel.

Discussion

Using partial proteolysis, I have shown that both p64 and p110 can be cleaved into distinct fragments. Because both proteins contain multiple sites that are not cleaved under my conditions, I interpret these results to mean that only some of the cleavage sites are accessible to proteases. The simplest explanation for this selective cleavage is that the amino acids that constitute the inaccessible sites are within tightly packed structures and that the accessible sites are found in sequences that connect these distinct domains.

In crystal structures of the zinc cluster proteins Gal4p and Ppr1p, the zinc cluster domain is tethered to the rest of the protein by a flexible “linker” region (Marmorstein et

al., 1992; Marmorstein and Harrison, 1994). It was therefore expected that the zinc cluster could be cleaved from p64. N-terminal microsequencing of a 51 kilodalton proteolytic fragment of p64 showed that it was cleaved approximately 40 amino acids downstream of the first cleavage site, suggesting that the “linker” region of p64 contains at least 40 residues (data not shown). The structure of the remaining region of p64 was not thoroughly investigated, but it is expected that this region will enable both dimerization of p64 and interaction with other CBF3 proteins. Consistent with this proposal, p64 contains a putative coiled-coil region in residues 369-390.

p110 appears to contain three distinguishable physical domains. N-terminal microsequencing of these fragments may verify the proposed domain organization of p110 and determine the orientation of the domains relative to the amino- and carboxyl-termini. In the “extended” CBF3-DNA complex, additional p110 protein contacts 33 bases of non-specific DNA sequence to the right of the region bound in the “core” complex. 33 bases corresponds to approximately 112 angstroms of B-form DNA, which is 70% of the length of a p110 dimer. If I equate length with mass, which is certainly an oversimplification, 70% of a p110 monomer corresponds to 81 kilodaltons, approximately the size of the major tryptic proteolytic fragment of p110. This calculation suggests that the 72 kilodalton domain stretches along the DNA while the 20 and 25 kilodalton regions function in p110 dimerization and to bind p110 to other CBF3 proteins. In support of this model, p110 contains two short putative coiled-coil regions that might fall into the 25 kilodalton domain if the 20 and 25 kilodalton domains are oriented at the carboxyl-terminus of p110.

Acknowledgements

I am grateful to Peter Buckel for providing the pUBS520 plasmid, and to Ken Kaplan for donating p64 purified from insect cells.

Figure A.1. p64 expressed in insect cells and bacteria is equally competent to participate in DNA binding. H6-tagged p64 was purified from either baculovirus-infected insect cells (circles) or *E. coli* (squares) and added to excess p23, p58, and p110 in a CDEIII DNA bandshift assay. The amount of probe shifted was quantified and plotted relative to the amount of probe shifted by a positive control.

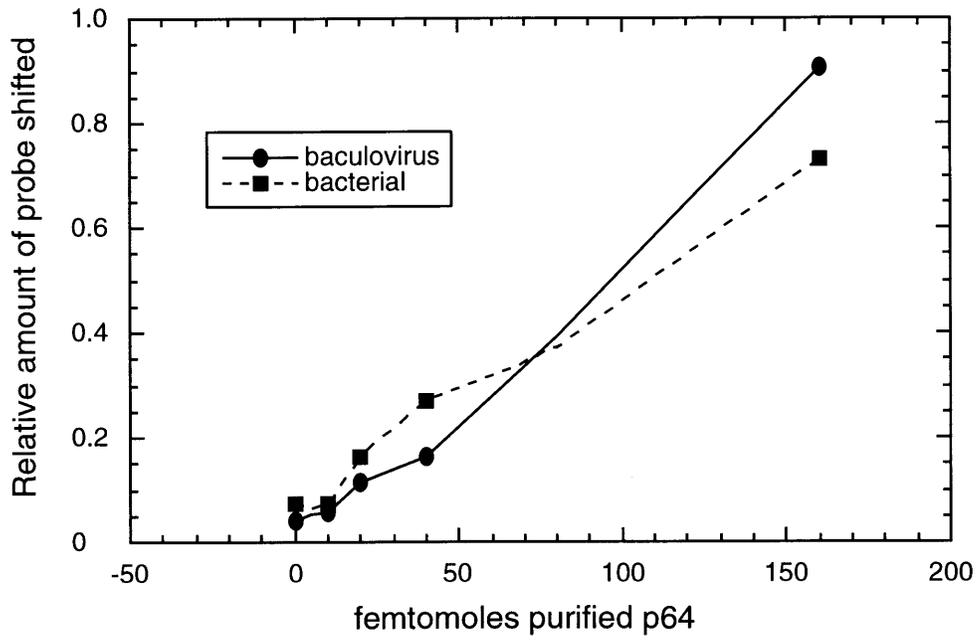
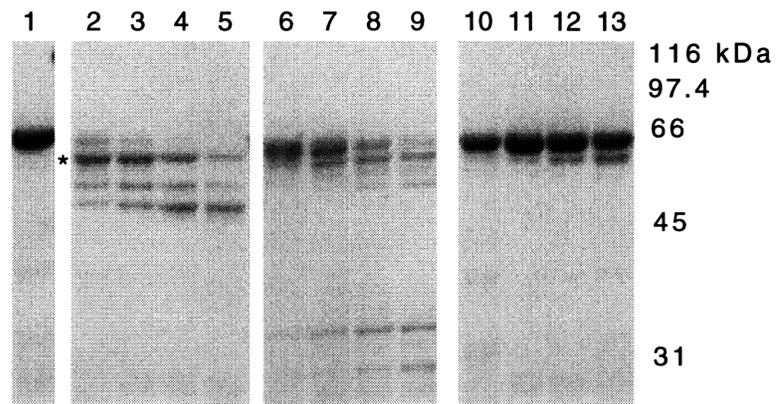


Figure A.2. Partial proteolysis of p64. (A) 0.2 mg/ml p64 purified from *E. coli* was treated with 5 μ g/ml trypsin (lanes 2-5), 1 μ g/ml chymotrypsin (lanes 6-9), or 2 μ g/ml V8 protease (lanes 10-13) at 30°. Aliquots of reactions were removed at 10 min (lanes 2, 6, 10), 20 min (lanes 3, 7, 11), 40 min (lanes 4, 8, 12), and 80 min (lanes 5, 9, 13) and stopped by boiling in SDS-PAGE loading buffer and PMSF. Samples were separated on an SDS-polyacrylamide gel and stained with Coomassie blue. Lane 1 contains undigested proteins. The tryptic fragment indicated with an asterisk was analyzed by N-terminal sequencing. (B) The sequence of the amino-terminal portion of p64 is shown. The arrow indicates the tryptic cleavage site, which is just downstream from the zinc cluster homology region (boxed).

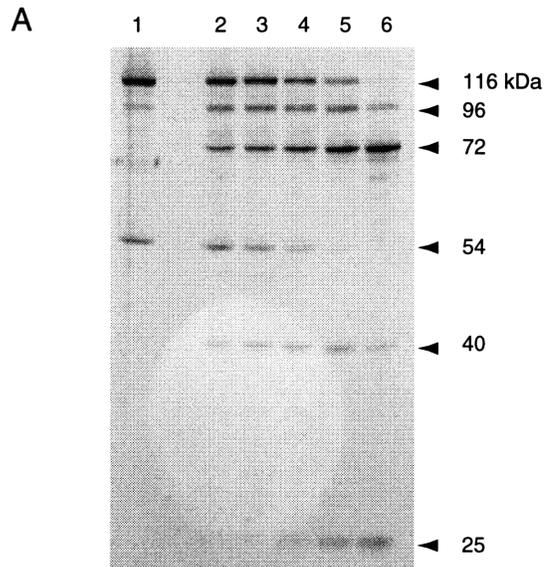
A



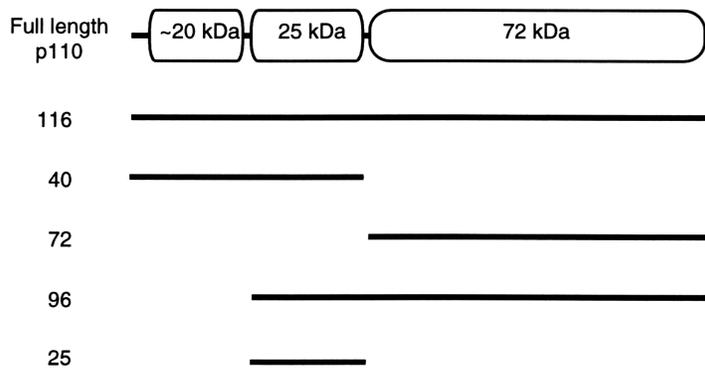
B

1 10 20 30 40 50
MFNRTTQLKSKHPCSVCTRRKVKCDRMI PCGNCRKRGQDSECMKSTKLITASSKEY
▲

Figure A.3. Partial proteolysis of p110. (A) 0.1 mg/ml purified p110 was digested with 0.5 μ g/ml trypsin. Aliquots were removed at 5, 10, 20, 40, and 80 minutes (lanes 2-6) and boiled in SDS-PAGE loading buffer with PMSF. Lane 1 contains undigested protein. The molecular weights of digestion products are indicated on the right. The 54 kilodalton protein is a contaminant in the p110 preparation that is recognized by antibodies directed against p110, suggesting it is a proteolytic degradation product produced during purification. (B) A model for the domain organization of p110 based on tryptic digestion. See text for a detailed description. The lines below the schematic of full length p110 indicate the different tryptic fragments used to create the model, with the molecular weights of the fragments indicated to the left.



B



Appendix B

Chromosome movement: Kinetochores motor along

This Appendix contains a reprint of the article, “Chromosome movement: Kinetochores motor along” by Adam Grancell and Peter K. Sorger, originally published in *Current Biology* 1998, **8**: R382-R385.

Chromosome movement: Kinetochores motor along

Adam Grancell and Peter K. Sorger

The equal division of chromosomes among daughter cells at mitosis involves a complex series of kinetochore-dependent chromosome movements. The kinetochore-associated CENP-E motor protein is critical for the sustained movement of chromosomes towards the metaphase plate during chromosome congression.

Address: Department of Biology, 68-371, MIT, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.
E-mail: asgrance@mit.edu; psorger@mit.edu

Current Biology 1998, 8:R382–R385
<http://biomednet.com/elecref/09609822008R0382>

© Current Biology Ltd ISSN 0960-9822

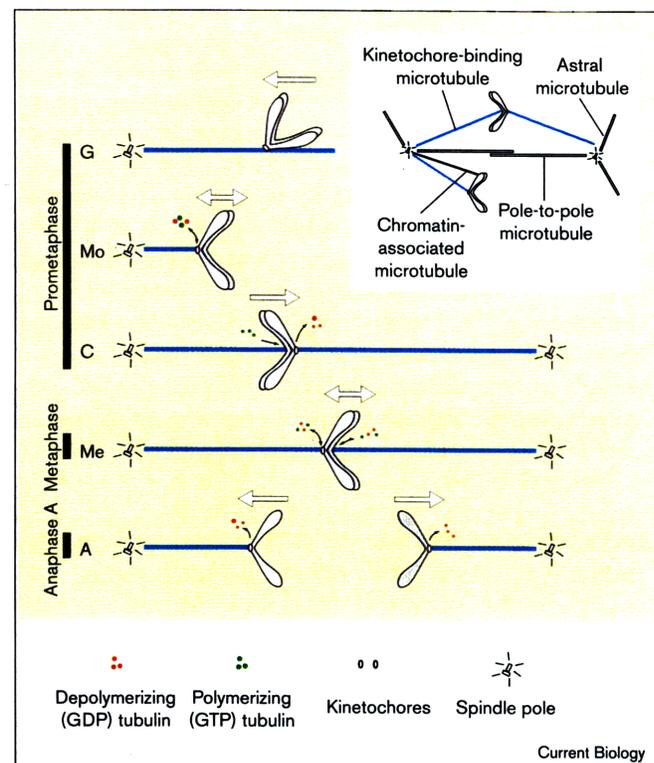
In eukaryotes, chromosome segregation occurs on a mitotic spindle formed from two polar arrays of microtubules. Chromosomes attach to the spindle microtubules by kinetochores, DNA–protein complexes that assemble on centromeric DNA. The kinetochore–microtubule linkage is very dynamic, and chromosomes move continuously during mitosis. Chromosome movement is powered by microtubule-based molecular motors, which use the energy of ATP hydrolysis to generate force, and by microtubule depolymerization, which liberates energy stored in the microtubule lattice by GTP hydrolysis during assembly [1]. A major challenge in the study of mitosis is to determine which kinetochore proteins are responsible for which aspects of chromosome movement, and to determine the relative importance of motors and microtubule dynamics in generating force. Several recent papers on the kinesin-related motor CENP-E advance our understanding of kinetochore-associated motors and illustrate the daunting technical challenges that must be overcome for chromosome movement to be understood in full.

To make sense of chromosome segregation, it is important to recall that spindle microtubules are oriented with their minus ends anchored at the spindle poles — the centrosomes — and their plus ends radiating outward. Spindle microtubules can be divided into four classes on the basis of the structures they contact (Figure 1, inset). The first class, astral microtubules, contact the cell cortex, where they play a role in orienting the spindle. The second class, pole-to-pole microtubules, overlap in the middle of the spindle with microtubules emanating from the opposite pole; motor proteins found in the overlap region cause pole-to-pole microtubules to slide relative to each other, thereby controlling the length of the spindle. The third class of microtubules interact with chromatin via chromokinesins, motor proteins on the chromosome arms which generate a polar ejection force that pushes

chromosomes toward the middle of the spindle. Lastly, the fourth class of microtubules bind to kinetochores to link chromosomes to the spindle. Although all four classes of microtubules play a role in positioning chromosomes, the force for chromosome movement is generated primarily by the kinetochore–microtubule attachment.

The purpose of mitosis is to separate newly replicated chromosomes into two equal and physically distinct sets. To ensure the accuracy of this process, all pairs of sister chromatids must achieve a state of bivalent attachment prior to their poleward movement at anaphase. Bivalent attachment is achieved when one of a pair of sister chromatids is

Figure 1



Kinetochore motility and microtubule assembly/disassembly in vertebrate mitotic cells. The kinetochore of a chromosome may initially become attached to the side of a microtubule and slide along its wall (G). Once the dynamic tips of the microtubules depolymerize and attach to a kinetochore, the monovalently-attached chromosome oscillates between slow movement towards, and slow movement away from, the pole (Mo). Once the sister kinetochore becomes attached to microtubules from the opposite pole, chromosomes congress at the spindle equator (C). When the chromosome is near the equator, at the location of the metaphase plate, each sister kinetochore alternates between movement towards and away from the pole (Me) until separation at anaphase (A). (Adapted from [1].)

attached via its kinetochore to microtubules emanating from one spindle pole, and the other sister is attached to microtubules from the opposite pole. Chromosomes make such bivalent attachments by a complex series of movements, as illustrated in Figure 1 [2].

Early in prometaphase, each pair of sister chromatids attaches by its kinetochore to the wall of a single microtubule, leading to rapid (20–50 μm per minute) poleward movement (Figure 1, state G). As the kinetochore-bound microtubule depolymerizes, wall-binding matures into an end-on attachment in which the extreme (plus) end of the microtubule is found at the kinetochore. A period of slow (2 μm per minute) movement, alternating away-from and towards the pole, then ensues (Figure 1, state Mo) until collision with the end of a microtubule emanating from the opposite pole produces bivalent attachment and slow (2 μm per minute), but sustained, movement towards the spindle equator (chromosome congression; Figure 1, state C). Pairs of sister chromatids then oscillate about the spindle equator (Figure 1, state Me) until the sudden loss of sister cohesion at anaphase allows movement of individual chromatids towards opposite poles (Figure 1, state A).

A remarkable feature of all but the earliest prometaphase chromosome–microtubule attachments is that the microtubule plus ends remain associated with kinetochores. This can occur because chromosome movement is tightly coupled to microtubule dynamics so that, as a chromosome moves, microtubules attached to the leading sister chromatid shrink and microtubules associated with the trailing sister grow. The mechanisms that couple movement to microtubule assembly are not known, but kinetochore-associated motor proteins are thought to play a key role in this coupling.

CENP-E structure and localization

With the aim of determining which proteins mediate chromosome movement, there has been a sustained attempt to identify kinetochore-associated microtubule-based motors. Thus far, cytoplasmic dynein [3,4] and two kinesin-like motor proteins — MCAK/XKCM1 [5,6] and CENP-E [7] — have been localized to mammalian kinetochores. We shall focus here on CENP-E, a 312 kDa protein with an amino-terminal kinesin-like motor domain, a carboxy-terminal microtubule-binding domain and a long intervening region that is likely to form a dimeric coiled-coil [8] (Figure 2, inset). One can imagine that a motor protein with such an architecture might link apposed microtubules. The microtubule-binding activity of CENP-E is turned on only at anaphase [8], however, implying that CENP-E can crosslink microtubules only late in mitosis.

Immuno-electron microscopy and light microscopy have been used to determine the subcellular distribution of

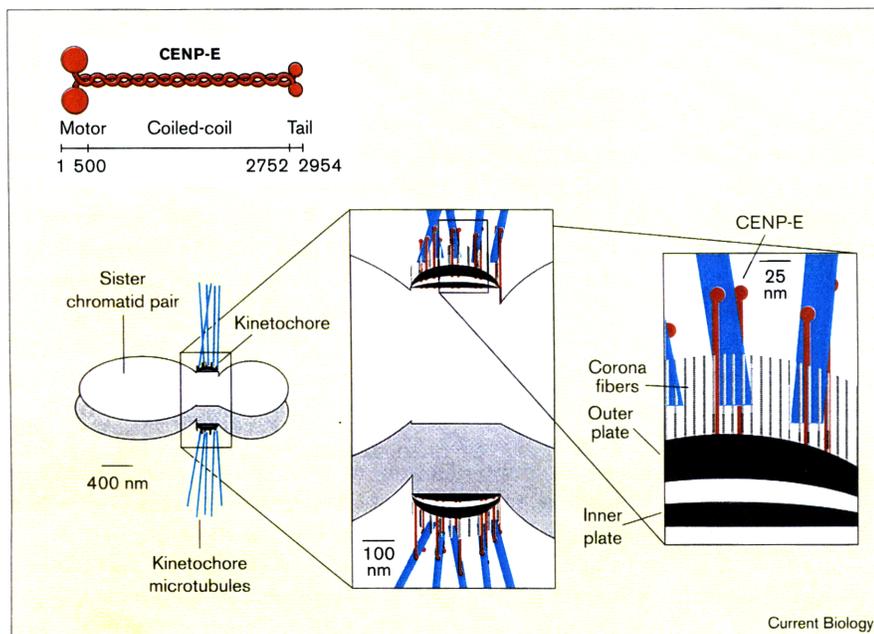
CENP-E during the cell cycle [9,10]. Early in prometaphase, just after nuclear envelope breakdown, most CENP-E is associated with microtubules. Subsequently, CENP-E becomes concentrated at kinetochores, presumably by traveling along microtubules by plus-end motor activity ([10] and see below). Kinetochore localization of CENP-E persists during metaphase and early anaphase. Late in anaphase, however, CENP-E dissociates from kinetochores, which are now near the spindle poles, and accumulates at the spindle midzone. The midzone contains the overlapping plus ends of pole-to-pole microtubules and is involved in the lengthening of the spindle in late anaphase (anaphase B). The relocalization of CENP-E to the midzone may reflect a role for CENP-E in crosslinking microtubules during spindle elongation.

Using immuno-electron microscopy, it was possible to show that CENP-E localizes to the interface between chromatin and microtubules even early in mitosis, when partially-assembled kinetochores have yet to assume mature electron-dense structures. It is reasonable to assume that these sites of CENP-E staining are the sites of nascent kinetochore formation [9,10]. In the subsequent period from prophase through anaphase, kinetochores are visible as distinct trilaminar structures and CENP-E is found in the fibrous region extending from the outer kinetochore plate and overlapping the ends of microtubules [9,10] (Figure 2). Because of its location, this fibrous corona is thought to contain the microtubule binding components of kinetochores.

CENP-E polarity

What is the polarity of the CENP-E motor? A motor's polarity is usually defined as the direction it moves on non-dynamic (usually taxol-stabilized) microtubules in the presence of ATP. Classical kinesin is a plus-end-directed motor, but some kinesin-like motors move towards the minus ends of microtubules; dynein is also a minus-end-directed motor. The oscillatory nature of chromosome movement implies a role for both plus-end-directed and minus-end-directed motor activity at kinetochores. Determining which types of movement are associated with which motors is, however, more tricky than it might appear. For example, in the absence of ATP, plus-end-directed kinesin can actually move cargo towards the minus ends of depolymerizing microtubules, apparently because the motors can hold on as depolymerization occurs (see below).

An early study [11] detected minus-end-directed microtubule motor activity in anti-CENP-E immunoprecipitates. More recently, however, Wood *et al.* [12] have shown that a bacterially-expressed fragment of CENP-E containing the motor domain has plus-end-directed motor activity *in vitro*. One possible explanation for this discrepancy is that the minus-end-directed motor activity brought

Figure 2

The structure (inset) and kinetochore localization (at three magnifications) of the CENP-E motor protein. The CENP-E stalk is potentially long enough for a single dimer to bind to any of the 20–30 microtubules that attach to a typical mammalian kinetochore. See text for details.

down by anti-CENP-E antibodies might have arisen from another motor protein associated with CENP-E. A second possibility is that CENP-E has both plus-end-directed and minus-end-directed motor activity, depending on its state of covalent modification or on its association with regulatory factors.

CENP-E function in the mitotic spindle

To investigate the functions of CENP-E in mitotic chromosome movement, Wood *et al.* [12] immunodepleted the motor from *Xenopus* oocyte extracts, and then examined spindle formation *in vitro*. In oocyte extracts from which CENP-E had been depleted, most sister chromatids failed to align at the metaphase plate. One possible explanation for this is that chromosomes skipped prometaphase and metaphase and entered anaphase prematurely. But sister chromatid separation — a marker for anaphase entry — had not occurred on the misaligned chromosomes, suggesting that CENP-E plays a role in either the movement of chromosomes towards the metaphase plate or maintaining them in that position once they are there.

Taking a different approach, Schaar *et al.* [13] used antibody microinjection to study CENP-E function in living cells. When injected into interphase cells, anti-CENP-E antibodies prevented metaphase chromosome alignment. In injected cells, some chromosomes formed only monopolar attachments and remained near one pole, and other chromosomes formed bipolar attachments but failed to congress. These results generally agree with those in *Xenopus* oocytes and show that, in living cells,

CENP-E is necessary to form bipolar spindle attachments and to move bivalently-attached chromosomes to the spindle equator.

Not all chromosome movement was abolished by anti-CENP-E antibodies, however. In injected cells, those chromosomes that had achieved bivalent attachment oscillated between the spindle poles, much like monovalently- and bivalently-attached metaphase chromosomes in unperturbed cells [2]. The existence of this oscillation in CENP-E-depleted cells shows that some aspects of chromosome movement were intact, even though the coordinated movements leading to congression were defective.

A model for CENP-E function

We can account for these observations by postulating that, during congression, the plus-end-directed motor activity of CENP-E is responsible for moving chromosomes along polymerizing microtubules towards their plus ends. It has been demonstrated, however, that pulling forces associated with the tips of shrinking microtubules (minus-end-directed movement), rather than pushing forces at the growing microtubule ends (plus-end-directed movement), provide the main force driving congression [14]. The failure of chromosomes to congress properly when CENP-E activity is compromised suggests that CENP-E plays a role in the minus-end-directed movement of chromosomes, even though it has a demonstrated plus-end-directed motor activity.

How could the plus-end-directed motor activity of CENP-E drive minus-end-directed movement? One possible

mechanism has been demonstrated in experiments by Lombillo *et al.* [15]. *In vitro*, chromosomes isolated from cultured cells bind via kinetochores to microtubule plus ends. When these microtubules are induced to depolymerize and motor movement is prevented by depleting ATP, the chromosomes remain attached to, and follow, the shrinking microtubule ends to generate minus-end-directed movement. This association with a shrinking microtubule is inhibited by adding antibodies against CENP-E. Thus it is possible that, in cells, CENP-E is involved in generating minus-end-directed movement by passively attaching chromosomes to shrinking microtubule ends. An alternative possibility, suggested by the observation that cells microinjected with anti-CENP-E antibodies still support oscillatory movement of chromosomes, is that CENP-E regulates movement towards the equator and is not required for the actual attachment of kinetochores to microtubule ends.

Conclusions

Kinetochores-mediated chromosome movement is dependent on a large number of different proteins, most of which have not yet been identified, that must act in concert to attach chromosomes to microtubules, control the oscillatory movements of prometaphase, maintain chromosome position at the metaphase plate, and move chromosomes poleward at anaphase. The task of studying this movement is confounded by the difficulty of separating the different types of motion *in vitro*, while maintaining a semblance of the motion that chromosomes undergo in living cells. Nevertheless, recent progress with CENP-E demonstrates that the identification and characterization of kinetochores-associated microtubule motor proteins is a good first step in the search for a mechanistic understanding chromosome segregation.

References

- Inoue S, Salmon ED: Force generation by microtubule assembly/disassembly in mitosis and related movements. *Mol Biol Cell* 1995, 6:1619-1640.
- Skibbens RV, Skeen VP, Salmon ED: Directional instability of kinetochores motility during chromosome congression and segregation in mitotic newt lung cells: a push-pull mechanism. *J Cell Biol* 1993, 122:859-875.
- Pfarr CM, Coue M, Grissom PM, Hays TS, Porter ME, McIntosh JR: Cytoplasmic dynein is localized to kinetochores during mitosis. *Nature* 1990, 345:263-265.
- Steuer ER, Wordeman L, Schroer TA, Sheetz MP: Localization of cytoplasmic dynein to mitotic spindles and kinetochores. *Nature* 1990, 345:266-268.
- Walczak CE, Mitchison TJ, Desai A: XKCM1: a *Xenopus* kinesin-related protein that regulates microtubule dynamics during mitotic spindle assembly. *Cell* 1996, 84:37-47.
- Wordeman L, Mitchison TJ: Identification and partial characterization of mitotic centromere-associated kinesin, a kinesin-related protein that associates with centromeres during mitosis. *J Cell Biol* 1995, 128:95-104.
- Yen TJ, Li G, Schaar BT, Szilak I, Cleveland DW: CENP-E is a putative kinetochores motor that accumulates just before mitosis. *Nature* 1992, 359:536-539.
- Liao H, Li G, Yen TJ: Mitotic regulation of microtubule cross-linking activity of CENP-E kinetochores protein. *Science* 1994, 265:394-398.
- Cooke CA, Schaar B, Yen TJ, Earnshaw WC: Localization of CENP-E in the fibrous corona and outer plate of mammalian kinetochores from prometaphase through anaphase. *Chromosoma* 1997, 106:446-455.
- Yao X, Anderson KL, Cleveland DW: The microtubule-dependent motor centromere-associated protein E (CENP-E) is an integral component of kinetochores corona fibers that link centromeres to spindle microtubules. *J Cell Biol* 1997, 139:435-447.
- Thrower DA, Jordan MA, Schaar BT, Yen TJ, Wilson L: Mitotic HeLa cells contain a CENP-E-associated minus end-directed microtubule motor. *EMBO J* 1995, 14:918-926.
- Wood KW, Sakowicz R, Goldstein LS, Cleveland DW: CENP-E is a plus end-directed kinetochores motor required for metaphase chromosome alignment. *Cell* 1997, 91:357-366.
- Schaar BT, Chan GK, Maddox P, Salmon ED, Yen TJ: CENP-E function at kinetochores is essential for chromosome alignment. *J Cell Biol* 1997, 139:1373-1382.
- Khodjakov A, Rieder CL: Kinetochores moving away from their associated pole do not exert a significant pushing force on the chromosome. *J Cell Biol* 1996, 135:315-327.
- Lombillo VA, Nislow C, Yen TJ, Gelfand VI, McIntosh JR: Antibodies to the kinesin motor domain and CENP-E inhibit microtubule depolymerization-dependent motion of chromosomes *in vitro* [see comments]. *J Cell Biol* 1995, 128:107-115.

Literature cited

- Allshire, R. C. (1997). Centromeres, checkpoints and chromatid cohesion. *Curr Opin Genet Dev* 7, 264-73.
- Baker, R. E., Fitzgerald-Hayes, M., and O'Brien, T.C. (1989). Purification of the yeast centromere binding protein CP1 and a mutational analysis of its binding site. *J Biol Chem* 264, 10843-50.
- Baker, R. E., and Masison, D. C. (1990). Isolation of the gene encoding the *Saccharomyces cerevisiae* centromere-binding protein CP1. *Mol Cell Biol* 10, 2458-67.
- Bascom-Slack, C. A., and Dawson, D. S. (1997). The yeast motor protein, Kar3p, is essential for meiosis I. *J Cell Biol* 139, 459-67.
- Basrai, M. A., Kingsbury, J., Koshland, D., Spencer, F., and Hieter, P. (1996). Faithful chromosome transmission requires Spt4p, a putative regulator of chromatin structure in *Saccharomyces cerevisiae*. *Mol Cell Biol* 16, 2838-47.
- Bell, S., and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* 357, 128-134.
- Bloom, K. S., and Carbon, J. (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell* 29, 305-17.
- Bram, R. J., and Kornberg, R. D. (1987). Isolation of a *Saccharomyces cerevisiae* centromere DNA-binding protein, its human homolog, and its possible role as a transcription factor. *Mol Cell Biol* 7, 403-9.
- Brinkmann, U., Mattes, R., and Buckel, P. (1989). High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product. *Gene* 85, 109-114.
- Brown, M. T., Goetsch, L., and Hartwell, L. H. (1993). *MIF2* is required for mitotic spindle integrity during anaphase spindle elongation in *Saccharomyces cerevisiae*. *J Cell Biol* 123, 387-403.
- Cahill, D. P., Lengauer, C., Yu, J., Riggins, G. J., Willson, J. K., Markowitz, S. D., Kinzler, K. W., and Vogelstein, B. (1998). Mutations of mitotic checkpoint genes in human cancers [see comments]. *Nature* 392, 300-3.
- Cai, M. J., and Davis, R. W. (1989). Purification of a yeast centromere-binding protein that is able to distinguish single base-pair mutations in its recognition site. *Mol Cell Biol* 9, 2544-50.

Cherry, J. M., Adler, C., Ball, C., Dwight, S., Chervitz, S., Juvik, G., Roe, T., Weng, S., and Botstein, D. "Saccharomyces Genome Database", <http://genome-www.stanford.edu/Saccharomyces/>.

Clarke, L., and Carbon, J. (1983). Genomic substitutions of centromeres in *Saccharomyces cerevisiae*. *Nature* 305, 23-8.

Clarke, L., and Carbon, J. (1980a). Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* 287, 504-9.

Clarke, L., and Carbon, J. (1980b). Isolation of the centromere-linked *CDC10* gene by complementation in yeast. *Proc Natl Acad Sci U S A* 77, 2173-7.

Connelly, C., and Hieter, P. (1996). Budding yeast *SKP1* encodes an evolutionarily conserved kinetochore protein required for cell cycle progression. *Cell* 86, 275-85.

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

Cottarel, G., Shero, J. H., Hieter, P., and Hegemann, J. H. (1989). A 125-base-pair *CEN6* DNA fragment is sufficient for complete meiotic and mitotic centromere functions in *Saccharomyces cerevisiae*. *Mol Cell Biol* 9, 3342-9.

Cumberledge, S., and Carbon, J. (1987). Mutational analysis of meiotic and mitotic centromere function in *Saccharomyces cerevisiae*. *Genetics* 117, 203-12.

Densmore, L., Payne, W. E., and Fitzgerald-Hayes, M. (1991). *In vivo* genomic footprint of a yeast centromere. *Mol Cell Biol* 11, 154-65.

Doheny, K. F., Sorger, P. K., Hyman, A. A., Tugendreich, S., Spencer, F., and Hieter, P. (1993). Identification of essential components of the *S. cerevisiae* kinetochore. *Cell* 73, 761-74.

Espelin, C. W., Kaplan, K. B., and Sorger, P. K. (1997). Probing the architecture of a simple kinetochore using DNA-protein crosslinking. *J Cell Biol* 139, 1383-96.

Feldman, R., Correll, C., Kaplan, K., and Deshaies, R. (1997). A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell* 91, 221-230.

Fitzgerald-Hayes, M., Buhler, J. M., Cooper, T. G., and Carbon, J. (1982). Isolation and subcloning analysis of functional centromere DNA (*CEN11*) from *Saccharomyces cerevisiae* chromosome XI. *Mol Cell Biol* 2, 82-7.

Fitzgerald-Hayes, M., Clarke, L., and Carbon, J. (1982). Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. *Cell* 29, 235-44.

Foreman, P. K., and Davis, R. W. (1993). Point mutations that separate the role of *Saccharomyces cerevisiae* centromere binding factor 1 in chromosome segregation from its role in transcriptional activation. *Genetics* 135, 287-96.

Funk, M., Hegemann, J., and Philippsen, P. (1989). Chromatin digestion with restriction endonuclease reveals 150-160 bp of protected DNA in the centromere of chromosome XIV in *Saccharomyces cerevisiae*. *Mol Gen Genet* 219, 153-160.

Gaudet, A., and Fitzgerald-Hayes, M. (1987). Alterations in the adenine-plus-thymine-rich region of *CEN3* affect centromere function in *Saccharomyces cerevisiae*. *Mol Cell Biol* 7, 68-75.

Goh, P. Y., and Kilmartin, J. V. (1993). *NDC10*: a gene involved in chromosome segregation in *Saccharomyces cerevisiae*. *J Cell Biol* 121, 503-12.

Hartwell, L., Dutcher, S., Wood, J., and Garvik, B. (1982). The fidelity of mitotic chromosome reproduction in *S. cerevisiae*. *Recent Adv. Yeast Mol. Biol.* 1, 28-38.

Hegemann, J. H., Shero, J. H., Cottarel, G., Philippsen, P., and Hieter, P. (1988). Mutational analysis of centromere DNA from chromosome VI of *Saccharomyces cerevisiae*. *Mol Cell Biol* 8, 2523-35.

Heus, J. J., Zonneveld, B. J., Steensma, H. Y., and Van den Berg, J. A. (1994). Mutational analysis of centromeric DNA elements of *Kluyveromyces lactis* and their role in determining the species specificity of the highly homologous centromeres from *K. lactis* and *Saccharomyces cerevisiae*. *Mol Gen Genet* 243, 325-33.

Hieter, P., Mann, C., Snyder, M., and Davis, R. W. (1985). Mitotic stability of yeast chromosomes: a colony color assay that measures nondisjunction and chromosome loss. *Cell* 40, 381-92.

Hieter, P., Pridmore, D., Hegemann, J. H., Thomas, M., Davis, R. W., and Philippsen, P. (1985). Functional selection and analysis of yeast centromeric DNA. *Cell* 42, 913-21.

Hoyt, M., He, L., Loo, K., and Saunders, W. (1992). Two *Saccharomyces cerevisiae* kinesin-related gene products required for mitotic spindle assembly. *J Cell Biol* 118, 109-120.

Hsiao, C. L., and Carbon, J. (1981). Direct selection procedure for the isolation of functional centromeric DNA. *Proc Natl Acad Sci U S A* 78, 3760-4.

Jaffe, M. (1996). Biochemical studies of the *Saccharomyces cerevisiae* kinetochore. Ph.D. Thesis (Massachusetts Institute of Technology, Cambridge, MA).

- Jehn, B., Niedenthal, R., and Hegemann, J. H. (1991). *In vivo* analysis of the *Saccharomyces cerevisiae* centromere CDEIII sequence: requirements for mitotic chromosome segregation. *Mol Cell Biol* *11*, 5212-21.
- Jiang, W., Lechner, J., and Carbon, J. (1993). Isolation and characterization of a gene (*CBF2*) specifying a protein component of the budding yeast kinetochore. *J Cell Biol* *121*, 513-9.
- Jiang, W., Lim, M. Y., Yoon, H. J., Thorner, J., Martin, G. S., and Carbon, J. (1995). Overexpression of the yeast *MCK1* protein kinase suppresses conditional mutations in centromere-binding protein genes *CBF2* and *CBF5*. *Mol Gen Genet* *246*, 360-6.
- Jiang, W., Middleton, K., Yoon, H. J., Fouquet, C., and Carbon, J. (1993). An essential yeast protein, CBF5p, binds *in vitro* to centromeres and microtubules. *Mol Cell Biol* *13*, 4884-93.
- Jiang, W. D., and Philippsen, P. (1989). Purification of a protein binding to the CDEI subregion of *Saccharomyces cerevisiae* centromere DNA. *Mol Cell Biol* *9*, 5585-93.
- Kaplan, K. B., Hyman, A. A., and Sorger, P. K. (1997). Regulating the yeast kinetochore by ubiquitin-dependent degradation and Skp1p-mediated phosphorylation. *Cell* *91*, 491-500.
- Keegan, L., Gill, G., and Ptashne, M. (1986). Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* *231*.
- Kent, N. A., Tsang, J. S., Crowther, D. J., and Mellor, J. (1994). Chromatin structure modulation in *Saccharomyces cerevisiae* by centromere and promoter factor 1. *Mol Cell Biol* *14*, 5229-41.
- Kozarov, E., van der Wel, H., Field, M., Gritzali, M., Brown, R. D., Jr., and West, C. M. (1995). Characterization of FP21, a cytosolic glycoprotein from *Dictyostelium*. *J Biol Chem* *270*, 3022-30.
- Krek, W. (1998). Proteolysis and the G1-S transition: the SCF connection. *Curr Opin Genet Dev* *8*, 36-42.
- Kuras, L., Barbey, R., and Thomas, D. (1997). Assembly of a bZIP-bHLH transcription activation complex: formation of the yeast Cbf1-Met4-Met28 complex is regulated through Met28 stimulation of Cbf1 DNA binding. *Embo J* *16*, 2441-51.
- Lafontaine, D. L. J., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M., and Tollervey, D. (1998). The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev* *12*, 527-37.

Lechner, J. (1994). A zinc finger protein, essential for chromosome segregation, constitutes a putative DNA binding subunit of the *Saccharomyces cerevisiae* kinetochore complex, Cbf3. *Embo J* 13, 5203-11.

Lechner, J., and Carbon, J. (1991). A 240 kd multisubunit protein complex, CBF3, is a major component of the budding yeast centromere. *Cell* 64, 717-25.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution [see comments]. *Nature* 389, 251-60.

Mann, C., and Davis, R. W. (1983). Instability of dicentric plasmids in yeast. *Proc Natl Acad Sci U S A* 80, 228-32.

Mann, C., and Davis, R. W. (1986). Structure and sequence of the centromeric DNA of chromosome 4 in *Saccharomyces cerevisiae*. *Mol Cell Biol* 6, 241-5.

Marmorstein, R., Carey, M., Ptashne, M., and Harrison, S. C. (1992). DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* 356, 408-414.

Marmorstein, R., and Harrison, S. C. (1994). Crystal structure of a PPR1-DNA complex: DNA recognition by proteins containing a Zn₂Cys binuclear cluster. *Genes Dev* 8, 2504-2512.

Mellor, J., Rathjen, J., Jiang, W., Barnes, C. A., and Dowell, S. J. (1991). DNA binding of CPF1 is required for optimal centromere function but not for maintaining methionine prototrophy in yeast [published erratum appears in *Nucleic Acids Res* 1991 Sep 25;19(18):5112]. *Nucleic Acids Res* 19, 2961-9.

Meluh, P., and Rose, M. (1990). *KAR3*, a kinesin-related gene required for yeast nuclear fusion. *Cell* 60, 1029-1041.

Meluh, P. B., and Koshland, D. (1997). Budding yeast centromere composition and assembly as revealed by *in vivo* cross-linking. *Genes Dev* 11, 3401-12.

Meluh, P. B., and Koshland, D. (1995). Evidence that the *MIF2* gene of *Saccharomyces cerevisiae* encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Mol Biol Cell* 6, 793-807.

Middleton, K., and Carbon, J. (1994). *KAR3*-encoded kinesin is a minus-end-directed motor that functions with centromere binding proteins (CBF3) on an *in vitro* yeast kinetochore. *Proc Natl Acad Sci U S A* 91, 7212-6.

Mortimer, R. K., and Schild, D. (1980). Genetic map of *Saccharomyces cerevisiae*. *Microbiol Rev* 44, 519-71.

- Murphy, M. R., Fowlkes, D. M., and Fitzgerald-Hayes, M. (1991). Analysis of centromere function in *Saccharomyces cerevisiae* using synthetic centromere mutants. *Chromosoma* 101, 189-97.
- Murphy, T., and Karpen, G. (1998). Centromeres take flight: Alpha satellite and the quest for the human centromere. *Cell* 93, 317-320.
- Murray, A. W., and Szostak, J. W. (1983). Pedigree analysis of plasmid segregation in yeast. *Cell* 34, 961-70.
- Nakayama, C., Teplow, D., and Harshey, R. (1987). Structural domains in phage Mu transposase: identification of the site-specific DNA-binding domain. *Proc Natl Acad Sci USA* 84, 1809-1813.
- Ng, R., and Carbon, J. (1987). Mutational and *in vitro* protein-binding studies on centromere DNA from *Saccharomyces cerevisiae*. *Mol Cell Biol* 7, 4522-34.
- Nicklas, R. B. (1997). How cells get the right chromosomes. *Science* 275, 632-7.
- Niedenthal, R., Stoll, R., and Hegemann, J. H. (1991). *In vivo* characterization of the *Saccharomyces cerevisiae* centromere DNA element I, a binding site for the helix-loop-helix protein CPF1. *Mol Cell Biol* 11, 3545-53.
- O'connell, K. F., Surdin-Kerjan, Y., and Baker, R. E. (1995) Role of the *Saccharomyces cerevisiae* general regulatory factor CP1 in methionine biosynthetic gene transcription. *Mol Cell Biol* 15, 1879-88.
- Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes Dev* 10, 2657-2683.
- Orr-Weaver, T. L., and Weinberg, R. A. (1998). A checkpoint on the road to cancer [news; comment]. *Nature* 392, 223-4.
- Panzeri, L., and Philippsen, P. (1982). Centromeric DNA from chromosome VI in *Saccharomyces cerevisiae* strains. *Embo J* 1, 1605-11.
- Peterson, J. B., and Ris, H. (1976). Electron-microscopic study of the spindle and chromosome movement in the yeast *Saccharomyces cerevisiae*. *J Cell Sci* 22, 219-42.
- Pfarr, C., Coue, M., Grissom, P., Hays, T., Porter, M., and McIntosh, J. (1990). Cytoplasmic dynein is localized to kinetochores during mitosis. *Nature* 345, 263-265.
- Pfeifer, K., Kim, K.-S., Kogan, S., and Guarente, L. (1989). Functional dissection and sequence of yeast *HAPI* activator. *Cell* 56, 291-301.

Reece, R., and Ptashne, M. (1993). Determinants of binding-site specificity among yeast C6 zinc cluster proteins. *Science* *261*, 909-911.

Reeves, R., and Nissen, M. (1990). The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J Biol Chem* *265*, 8573-8582.

Saitoh, H., Tomkiel, H., Cooke, C., Ratrie, H. d., Maurer, M., Rothfield, N., and Earnshaw, W. (1992). CENP-C, an autoantigen in sclerodoma, is a component of the human inner kinetochore plate. *Cell* *70*, 115-125.

Saunders, M., Fitzgerald-Hayes, M., and Bloom, K. (1988). Chromatin structure of altered yeast centromeres. *Proc Natl Acad Sci U S A* *85*, 175-9.

Saunders, M. J., Yeh, E., Grunstein, M., and Bloom, K. (1990). Nucleosome depletion alters the chromatin structure of *Saccharomyces cerevisiae* centromeres. *Mol Cell Biol* *10*, 5721-7.

Schaar, B. T., Chan, G. K., Maddox, P., Salmon, E. D., and Yen, T. J. (1997). CENP-E function at kinetochores is essential for chromosome alignment. *J Cell Biol* *139*, 1373-82.

Sears, D. D., Hegemann, J. H., Shero, J. H., and Hieter, P. (1995). *Cis*-acting determinants affecting centromere function, sister-chromatid cohesion and reciprocal recombination during meiosis in *Saccharomyces cerevisiae*. *Genetics* *139*, 1159-73.

Siegel, L. M., and Monty, K. J. (1966). Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases. *Biochim. Biophys. Acta.* *112*, 346-362.

Skowyra, D., Craig, K., Tyers, M., Elledge, S., and Harper, J. (1997). F-box proteins are receptors that recruit phosphorylated substrates to SCF ubiquitin-ligase complex. *Cell* *91*, 209-219.

Smith, M. M., Yang, P., Santisteban, M. S., Boone, P. W., Goldstein, A. T., and Megee, P. C. (1996). A novel histone H4 mutant defective in nuclear division and mitotic chromosome transmission. *Mol Cell Biol* *16*, 1017-26.

Sober, H. A. (1970). *Handbook of Biochemistry* (Cleveland: Chemical Rubber Co.).

Sorger, P. K., Doheny, K. F., Hieter, P., Kopski, K. M., Huffaker, T. C., and Hyman, A. A. (1995). Two genes required for the binding of an essential *Saccharomyces cerevisiae* kinetochore complex to DNA. *Proc Natl Acad Sci U S A* *92*, 12026-30.

Sorger, P. K., Severin, F. F., and Hyman, A. A. (1994). Factors required for the binding of reassembled yeast kinetochores to microtubules *in vitro*. *J Cell Biol* *127*, 995-1008.

- Stemmann, O., and Lechner, J. (1996). The *Saccharomyces cerevisiae* kinetochore contains a cyclin-CDK complexing homologue, as identified by *in vitro* reconstitution. *Embo J* 15, 3611-20.
- Steuer, E., Wordeman, L., Schroer, T., and Sheetz, M. (1990). Localization of cytoplasmic dynein to mitotic spindles and kinetochores. *Nature* 345, 266-268.
- Stinchcomb, D. T., Mann, C., and Davis, R. W. (1982). Centromeric DNA from *Saccharomyces cerevisiae*. *J Mol Biol* 158, 157-90.
- Stoler, S., Keith, K. C., Curnick, K. E., and Fitzgerald-Hayes, M. (1995). A mutation in *CSE4*, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. *Genes Dev* 9, 573-86.
- Straight, A. F. (1997). Cell cycle: checkpoint proteins and kinetochores. *Curr Biol* 7, R613-6.
- Strunnikov, A. V., Kingsbury, J., and Koshland, D. (1995). *CEP3* encodes a centromere protein of *Saccharomyces cerevisiae*. *J Cell Biol* 128, 749-60.
- Walczak, C., Mitchison, T., and Desai, A. (1996). XKCM1: a *Xenopus* kinesin-related protein that regulates microtubule dynamics during mitotic spindle assembly. *Cell* 84, 37-47.
- Warburton, P. E., Cooke, C. A., Bourassa, S., Vafa, O., Sullivan, B. A., Stetten, G., Gimelli, G., Warburton, D., Tyler-Smith, C., Sullivan, K. F., Poirier, G. G., and Earnshaw, W. C. (1997). Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. *Curr Biol* 7, 901-4.
- Wiens, G., and Sorger, P. (1998). Centromeric chromatin and epigenetic effects in kinetochore assembly. *Cell* 93, 313-316.
- Wilmen, A., Pick, H., Niedenthal, R. K., Sen-Gupta, M., and Hegemann, J. H. (1994). The yeast centromere CDEI/Cpf1 complex: differences between *in vitro* binding and *in vivo* function. *Nucleic Acids Res* 22, 2791-800.
- Winey, M., Mamay, C. L., ET, O. T., Mastronarde, D. N., Giddings, T. H., Jr., McDonald, K. L., and McIntosh, J. R. (1995). Three-dimensional ultrastructural analysis of the *Saccharomyces cerevisiae* mitotic spindle. *J Cell Biol* 129, 1601-15.
- Wood, K. W., Sakowicz, R., Goldstein, L. S., and Cleveland, D. W. (1997). CENP-E is a plus end-directed kinetochore motor required for metaphase chromosome alignment. *Cell* 91, 357-66.

Wordeman, L., and Mitchison, T. (1995). Identification and partial characterization of mitotic centromere-associated kinesin, a kinesin-related protein that associates with centromeres during mitosis. *J Cell Biol* 128, 95-104.

Yao, X., Anderson, K. L., and Cleveland, D. W. (1997). The microtubule-dependent motor centromere-associated protein E (CENP- E) is an integral component of kinetochore corona fibers that link centromeres to spindle microtubules. *J Cell Biol* 139, 435-47.

Yen, T., Li, G., Schaar, B., Szilak, I., and Cleveland, D. (1992). CENP-E is a putative kinetochore motor that accumulates just before mitosis. *Nature* 359.

Zhang, L., and Guarente, L. (1994). The yeast activator *HAPI*--a *GAL4* family member--binds DNA in a directly repeated orientation. *Genes Dev* 8, 2210-2119.